

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2003, 13:01:09 ; Search time 2756 Seconds
(without alignments)
10983.079 Million cell updates/sec

Title: US-09-966-803-1

Perfect score: 1869
Sequence: 1 ATGACCGGCAATCGAATGGA.....TCAGGGCAATGGCACTGA 1869

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estnu:*
4: em_estov:*
5: em_estov:*
6: em_estov:*
7: em_estov:*
8: em_estov:*
9: em_estov:*
10: em_estov:*
11: em_estov:*
12: em_estov:*
13: em_estov:*
14: em_estov:*
15: em_estov:*
16: em_estov:*
17: em_estov:*
18: em_gss_hum:*
19: em_gss_hum:*
20: em_gss_hum:*
21: em_gss_hum:*
22: em_gss_hum:*
23: em_gss_hum:*
24: em_gss_hum:*
25: em_gss_hum:*
26: em_gss_hum:*
27: em_gss_hum:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	138.2	7.4	327	9	AA071581	AA071581 0028F PYR
2	53.2	2.8	803	13	BM321042	BM321042 rockefell
3	52.2	2.8	664	13	BM645896	BM645896 17006873
4	49.6	2.7	702	13	BM321397	BM321397 rockefell
5	49.2	2.6	453	13	AI323077	AI323077 m64c10.y
6	48.6	2.6	478	12	BG228520	BG228520 ux44g06.x

7	48	2.6	1341	11	AY108285	AY108285 zea_mays
8	47.8	2.6	870	12	BE906698	BE906698 601500641
9	47.8	2.6	967	9	AL528820	AL528820 AL528820
10	47.4	2.5	1227	13	BM548451	BM548451 AGENCOURT
11	47	2.5	464	12	BF727266	BF727266 IL5-17002
12	47	2.5	523	9	AI323432	AI323432 m64c10.x
13	47	2.5	548	12	BF721977	BF721977 mab26405
14	47	2.5	614	10	BE585051	BE585051 7-7A-20.F
15	47	2.5	958	14	AL043279	AL043279 DKF2P341
16	47	2.5	1329	11	AK011861	AK011861 Mus muscu
17	47	2.5	1330	11	AK011860	AK011860 Mus muscu
18	47	2.5	517	10	BE402639	BE402639 CSB010A09
19	46.6	2.5	517	10	BE402638	BE402638 BRT_4272
20	46.6	2.5	539	14	BQ608368	BQ608368 BRT_4272
21	46.4	2.5	564	14	BM929321	BM929321 UI-E-BJ1-
22	46.4	2.5	768	17	BF057277	BF057277 7R18e08.x
23	46.4	2.5	789	9	AL529961	AL529961 Tetradon
24	46.4	2.5	809	13	BE585057	BE585057 603028971
25	46.4	2.5	676	10	BE584073	BE584073 3-5E-HA.F
26	46	2.5	736	14	BQ806284	BQ806284 WHE3577A
27	46	2.5	755	14	BQ802857	BQ802857 WHE2830_G
28	46	2.5	1141	11	AY103772	AY103772 zea_mays
29	45.8	2.5	567	14	BO606563	BO606563 BRX_2428
30	45.6	2.4	343	10	AM802212	AM802212 IL5-UM007
31	45.6	2.4	750	10	BE122887	BE122887 16.05_Hum
32	45.6	2.4	852	13	BI752028	BI752028 60302143
33	45.6	2.4	458	14	BM710892	BM710892 UI-E-DX1-
34	45.4	2.4	527	12	BG228504	BG228504 ux44d07.x
35	45.4	2.4	567	9	AL504599	AL504599 AL504599
36	45.4	2.4	926	14	BQ222794	BQ222794 AGENCOURT
37	45.2	2.4	582	9	AT734603	AT734603 606033605
38	45.2	2.4	682	14	BO806138	BO806138 WHE3575_C
39	45.2	2.4	609	13	BQ281768	BQ281768 BQ281768
40	45	2.4	568	13	BM354101	BM354101 IF30d06.Y
41	44.8	2.4	357	10	AM597116	AM597116 sj171912.Y
42	44.6	2.4	826	12	BF342052	BF342052 602012804
43	44.6	2.4	756	14	BO839085	BO839085 WHE3589.B
44	44.4	2.4	826	12	BF624680	BF624680 HYSME001
45	44.4	2.4				

ALIGNMENTS

RESULT 1
AA071581/c 327 bp mRNA linear EST 01-OCT-1996
LOCUS 0028F Pyrococcus furiosus 1-2AP II library, F Robb Pyrococcus
DEFINITION furiosus cDNA clone 0028 similar to Acylamine Acid releasing enzyme
mRNA sequence.

ACCESSION AA071581 GI:1578969
VERSION AA071581.1
KEYWORDS Pyrococcus furiosus.
SOURCE Pyrococcus furiosus.
ORGANISM Archaea; Euryarchaeota; Thermococci; Thermococcales;

REFERENCE 1 (bases 1 to 327)
Borges,K.M., Brummet,S.R., Bogert,A., Davis,M.C., Hujer,K.M., Domke
S.T., Szasz,J., Ravel,J., Diruggiero,J., Fuller,C., Chase,J.W. and
Robb,F.T.
A Survey of the Genome of the Hyperthermophilic Archaeon,
Pyrococcus furiosus

JOURNAL
COMMENT Unpublished (1996)
Contact: Brummet SR, Robb F
Brummet, Sequencing R&D
Amersham Life Science, Inc.
Amersham Life Science, 26101 Miles Rd, Cleveland, OH 44128

FEATURES
Seq primer: 17.
Location/Qualifiers

Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source
1. 967

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDD001YF08"
/clone_1ib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 146 a 326 c 268 g 218 t 9 others

ORIGIN

Query Match 2.6%; Score 47.8; DB 9; Length 967;
Best Local Similarity 49.0%; Pred. No. 0.67;

Matches 121; Conservative 2; Mismatches 124; Indels 0; Gaps 0;

1062 CTTCGGCTTCAGAGCAAGACCTCGATAGAGCGCTGTACTCAGCCGAGGTAA 1121
118 CATTGCCCTTGAATTTGGCCCTTACCTGAGAGTCTTGCCAGACCTTCATCAGCGTCTT 177
1122 AGAGGAGAGAGCCCGGATAGTCTTCTGTCACAGCGCGGCGAAGGCGATAGAGACA 1181
178 TGGCCACCTGTGTCTCTATGCTCAGCTCCGCTCTCTATGTTTACCTGCTCTA 237
1182 CGCTTCGTCTACAGAGATCAGCTGATGCGAGCAAGGCGCTACTGCTGCTTCTGAA 1241
238 TCTCTCTTCCGATGGACAGCTGAGAAATTTCAGAGGCGACCTACTGCTACTTGTGCC 297
1242 CCCGCGCGGACGAGCGCTATAGCAGACTTCCGCTCCGCTCTGAGAGAGACTGG 1301
298 CTACCTGGTGTCTTATGTGTGTGAGCATCCGCTGTCATCTCTGTGAGATTCACCG 357
1302 CTTGGAG 1308
358 CCTGGGG 364

RESULT 10

BM548451

LOCUS 1227 bp mRNA linear EST 20-FEB-2002

DEFINITION AGENCOURT_5573227 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5731944

ACCESSION BM548451

VERSION BM548451.1 GI:18783010

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE NIH-MGC http://mhc.nci.nih.gov/

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Invitrogen

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1273 row: e column: 01
High quality sequence stop: 487.

FEATURES

source
1. 1227

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5731944"
/clone_1ib="NIH_MGC_124"
/tissue_type="hippocampus"
/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; site:1: EcoRV (destroyed); site:2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb. Insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."

BASE COUNT 239 a 476 c 275 g 237 t

ORIGIN

Query Match 2.5%; Score 47.4; DB 13; Length 1227;
Best Local Similarity 46.7%; Pred. No. 0.94;

Matches 150; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

1062 CTTCGGCTTCAGAGCAAGACCTCGATAGAGCGCTGTACTCAGCCGAGGTAA 1121
174 CATTGCCCTTGAATTTGGCCCTTACCTGAGAGTCTTGCCAGACCTTCATCAGCGTCTT 233
1122 AGAGGAGAGAGCCCGGATAGTCTTCTGTCACAGCGCGGCGAAGGCGATAGAGACA 1181
234 CGGCCACCTGTGTCTCTATGCTCAGCTCCGCTCTCTATGTTTACCTGCTCTA 293
1182 CGCTTCGTCTACAGAGATCAGCTGATGCGAGCAAGGCGCTACTGCTGCTGAA 1241
294 TCTCTCTTCCGATGGACAGCTGAGAAATTTCAGAGGCGACCTACTGCTTGTGCC 353
1242 CCCGCGCGGACGAGCGCTATAGCAGACTTCCGCTCCGCTCTGAGAGAGACTGG 1301
354 CTACCTGGTGTCTTATGTGTGTGAGCTCCGCTGTCATCTGTGAGATTCACCG 413
1302 CTGAGAGACTTGTGAGACATATGACGAGCTGAGAGTCTTCAAGCTGAGACCGA 1361
414 CCGGGGCGTGTCCGCGCTCCATCGGCTCTCTCTCTTGTGCCCTCCATCT 473
1362 GGCGGACAGGAGCGGCTGG 1382
474 GGCGGCGGCGCTGTGCCCTTGG 494

RESULT 11

BF772766

LOCUS 464 bp mRNA linear EST 12-JAN-2001

DEFINITION IL5-IT0027-131200-330-f12 IT0027 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF772766

VERSION BF772766.1 GI:12120666

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Brito, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,

Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

Db	316	CTCGAGTCCCCACGGAGAAACCTCAACGAGCCCTGCAGCGCCATTAAAGCGACGTACACG	375
QY	1711	GCAGTATTCAAAGCCGGCGGCCACAGGCCACACAGCTCGCGGAGAGCCGACAGGCCG	1770
Db	376	CCCGAGCGTCCCTCGTCGCAAGCGGAGAGCAGCGCTTCCGACGCGTGGGTGCATCTCG	435
QY	1771	AAGGCGTACAGGCGTTCATGAGTTCTTCGAGGCGAAGCTCAGAAAGTACGAGAGAGGC	1833
Db	436	GGCCGCAAGGTGCTTCGAAAGCCCGCTACAAAGAGGCGGTGCGGTATGCTGCTGAAGGC	495
QY	1831	TTTGAGGTAGAGAAG	1845
Db	496	GTCGAGTAGAGACG	510

RESULT 15
AT 043270

LOCUS	AL043379	751 bp	mRNA	linear	EST 29-FEB-2000
DEFINITION	DKR2p33410423_r1 434 (synonym: htes3) Homo sapiens cDNA clone				
ACCESSION	DKR2p33410423	5'	mRNA sequence.		
VERSION	AL043379				
KEYWORDS	AL043279.2	GI:5935855			
SOURCE	EST.				
ORGANISM	human.				

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryotes; Metazoa; Chordata; Crustacea; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.
1 (Pages 1 to 751)
Blum, H., Baerachsch, S., Mewes, H.W., Gassenhuber, J., and Wiemann, S
EST (Blum, et al.)
Unpublished, (1999)
On Jul 8, 1999 this sequence version replaced g1:5422672.

On Jul 8, 1999 this sequence version replaced gl:5422672
Contact: Blum H
MIPS

Am Klopperspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by LMU (Ludwig Maximilians Universität,
Munich/Germany) within the CDNA sequencing consortium of the German
Genome Project.

This clone (DKFZp434L0423) is available at the RZPD in Berlin. No SI sequence available.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14055 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Source

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="DKFZp434L0423"  
/clone_lib="434 (synonym: htcs3)"  
/tissue_type="testis"  
/dev_stage="adult"
```

BASE COUNT	/note-Vector: pSPORT1; Site_1: NotI; Site_2: SalI*			
ORIGIN	108 a	264 c	181 g	191 t
				7 others

Query Match	2.58;	Score 47;	DB 9;	Length 751;
Best Local Similarity	48.58;	Pred. No. 0.93;		
Matches 128; Conservative	0;	Mismatches 136;	Indels 0;	Gaps 0;

QY	1062	CTTC	CCGGTTCAAGAGCAAAAGAC	CCGCGAGATAGACGGCGTGATACCTCAGACGGCGAGGTAA	1122
Db	391	CAATGCCCTTGATATTGGCCCTACCTCAGAAAGGTCCTTTGGCCAGACCTTCAACACCGTCCGT	450		
QY	1122	AGAGGAGAAGGCCCCGGTGATAGTCTTGCTGCACAGCGCGCGCGAAGGGCATGTACGACA	1188		
Db	451	CGGCGACCTGTGCTCCTCAATGTGAGCGTCCCGCTCTGTATGTCTACCTGCTCTA	510		
QY	1182	CCGCTCGTGTACAGATGCAAGTGCATGATGGCGAGCAAGGCGTACTACCTGCTCGTGA	1241		
Db	511	TCCTCTTTTTCGACATGCGACAGCTGAGGAAATTTCAAGGGCACTACCTCTTACCTTTGTCC	570		

QY	1242	CCCCGGGGGACGACAGGGCTCTTATACGAAGACTTGTGGCGTCCGCGCTCGGAGAGAGACTGG	1301
Db	571	CTACCTGCTGTCCTTACATGTCGTGTGAGCCTCTCGTGCTCAATCCGTCGTGGATCCAGCG	630
QY	1302	CTTGGAGAGACTTTGAGGACATAAT	1325
Db	631	CTGTGGGCTGCTNCGCGCTTCAT	654

Search completed: July 15, 2003, 16:11:46
Job time : 2761 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2003, 14:01:35 ; Search time 319 Seconds
(without alignments)
9251.230 Million cell updates/sec

Title: US-09-966-803-1
1869
Perfect score: 1 ATGACGGCATTCGATGGA.....TCAGGGGAATGGAACGTA 1869
Sequence: IDENTIFY_NUC
Scoring table: Gapped 10.0, Gapped 1.0

Searched: 1105431 seqs, 789497651 residues
Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/PCCTUS_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/PCCTUS_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1869	100.0	1869	10	US-09-966-803-1
2	140.8	7.5	1958	10	US-09-974-300-224
3	66.2	3.5	477	10	US-09-974-300-4866
4	55	2.9	783	10	US-09-815-242-8012
5	52.2	2.8	2817	9	US-10-156-761-5055
6	52.2	2.8	9025608	9	US-10-156-761-1
7	50	2.7	1446	9	US-10-156-761-2793
8	50	2.7	9025608	9	US-10-156-761-1
9	48.2	2.6	1503	9	US-10-156-761-6210
10	47	2.5	1785	9	US-09-864-761-18040
11	47	2.5	2428	10	US-09-822-849A-6
12	47	2.5	2673	9	US-10-060-425-1
13	46.6	2.5	18438	9	US-10-156-761-2673
14	46.6	2.5	125746	9	US-10-156-761-1886
15	46.6	2.5	125746	9	US-10-156-761-15102
16	45.8	2.5	3372	9	US-10-067-457-2
17	45.8	2.5	504	9	US-09-918-995-26097
18	45.6	2.4	1792	9	US-09-086-436-40
19	45	2.4	999	9	US-09-925-664-46

20	44.6	2.4	489	10	US-09-864-761-1278	Sequence 1278, Ap
21	44.4	2.4	594	9	US-10-123-155-10	Sequence 10, Appl
22	44.4	2.4	924	9	US-10-156-761-3907	Sequence 3907, Ap
23	43.8	2.3	1941	9	US-10-156-761-3418	Sequence 3418, Ap
24	43.4	2.3	681	9	US-10-156-761-3307	Sequence 3307, Ap
25	43.2	2.3	1155	9	US-10-156-761-4992	Sequence 4992, Ap
26	43	2.3	1556	9	US-10-081-872-121	Sequence 121, Appl
27	42.8	2.3	1176	9	US-10-156-761-3069	Sequence 3069, Ap
28	42.8	2.3	1299	9	US-09-977-719-180	Sequence 180, Appl
29	42.8	2.3	1809	9	US-10-156-761-4041	Sequence 4041, Ap
30	42.8	2.3	1899	9	US-10-156-761-7216	Sequence 7216, Ap
31	42.8	2.3	2051	9	US-09-975-719-114	Sequence 114, Appl
32	42.8	2.3	3012	9	US-10-156-761-7256	Sequence 7256, Ap
33	42.6	2.3	1272	9	US-10-156-761-2079	Sequence 2079, Ap
34	42.6	2.3	2149	10	US-10-171-581-103	Sequence 103, Appl
35	42.6	2.3	2149	10	US-09-880-107-3031	Sequence 3031, Ap
36	42.4	2.3	1014	9	US-10-156-761-5381	Sequence 5381, Ap
37	42.4	2.3	1512	10	US-09-880-107-3452	Sequence 3452, Ap
38	42.4	2.3	1632	9	US-10-156-761-6045	Sequence 6045, Ap
39	42	2.2	675	9	US-10-156-761-2995	Sequence 2995, Ap
40	42	2.2	1164	9	US-10-156-761-2298	Sequence 2298, Ap
41	42	2.2	1830	9	US-10-081-872-181	Sequence 181, Appl
42	41.8	2.2	1362	9	US-10-156-761-6703	Sequence 6703, Ap
43	41.6	2.2	1047	9	US-10-156-761-1815	Sequence 1815, Ap
44	41.4	2.2	1083	9	US-10-156-761-1132	Sequence 1132, Ap
45	41.2	2.2	1818	10	US-09-815-242-7989	Sequence 7989, Ap

ALIGNMENTS

RESULT 1
US-09-966-803-1
Sequence 1, Application US/09966803
Patent No. US20020137185A1
GENERAL INFORMATION:
APPLICANT: Murphy et al.
TITLE OF INVENTION: Amidases
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: CABELLA, BYRNE, BAIN, GILFILLAN,
CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/966, 803
FILING DATE: 27-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/609, 570
FILING DATE: 30-Jun-2000
APPLICATION NUMBER: 09/427, 372
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Charles J. Herion
REGISTRATION NUMBER: 28, 019
REFERENCE/DOCKET NUMBER: 331400-53
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1869 NUCLEOTIDES
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE

TOPOLOGY: LINEAR
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-966-803-1

Query Match 100.0%; Score 1869; DB 10; Length 1869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGACCGGCATCGAATGGAACACGACACCTTTCTAAGTTCGCTACCTGGCGACCCG 60
DB 1 ATGACCGGCATCGAATGGAACACGACACCTTTCTAAGTTCGCTACCTGGCGACCCG 60
QY 61 AGGATACGGGAACTTAATGCGGTACACCTGACGAGGCGCAACATGAAGAGACAAAG 120
DB 61 AGGATACGGGAACTTAATGCGGTACACCTGACGAGGCGCAACATGAAGAGACAAAG 120
QY 121 TACGAGACACGAGTTGTTGTAAGACCTTGAAGGCGCTCAAGCGCTTCATCGAGAC 180
DB 121 TACGAGACACGAGTTGTTGTAAGACCTTGAAGGCGCTCAAGCGCTTCATCGAGAC 180
QY 181 GCCCTCATGCGGAGGATTTGCGCGACGAGGCGAGAAAGCTCGCTTCACCTGCTTAACGAG 240
DB 181 GCCCTCATGCGGAGGATTTGCGCGACGAGGCGAGAAAGCTCGCTTCACCTGCTTAACGAG 240
QY 241 GAGAGAAGAGAGACGAGATATGGGTGGCGGATATCCAGACCCCTGAGCGCCCAAGAAAGTC 300
DB 241 GAGAGAAGAGAGACGAGATATGGGTGGCGGATATCCAGACCCCTGAGCGCCCAAGAAAGTC 300
QY 301 CTCTCAACTAATAAAGCTCCGCTCGATCGATGAGAACGAGATTCAGAGAGACTCTTAGTT 360
DB 301 CTCTCAACTAATAAAGCTCCGCTCGATCGATGAGAACGAGATTCAGAGAGACTCTTAGTT 360
QY 361 GTGGGCTTCAAGAGAGAGAGAGAGAGACTCTCTTTGACGACGAGCTCCGCTG 420
DB 361 GTGGGCTTCAAGAGAGAGAGAGAGAGACTCTCTTTGACGACGAGCTCCGCTG 420
QY 421 TTGACACATATGAGATCTTTGATGAGAGAGAGAGAGAGCTTGGGTTCTTGAAGTCTAG 480
DB 421 TTGACACATATGAGATCTTTGATGAGAGAGAGAGAGAGCTTGGGTTCTTGAAGTCTAG 480
QY 481 GCCGAGAGAGATATTCGACAGCTTGCAGAAAGCTTTGAGTTCGCTTCGACAGGCG 540
DB 481 GCCGAGAGAGATATTCGACAGCTTGCAGAAAGCTTTGAGTTCGCTTCGACAGGCG 540
QY 541 GATCGATAGTTGTTGAAGCTCCGCGACCGGAGGAGGAGCAAGCTTGCCTTGAAGTT 600
DB 541 GATCGATAGTTGTTGAAGCTCCGCGACCGGAGGAGGAGCAAGCTTGCCTTGAAGTT 600
QY 601 TACGACATATGCTATGGAAGAGAGGAGGAGAGAGAGAGCTTTCGAGAGGCTCTCTTC 660
DB 601 TACGACATATGCTATGGAAGAGAGGAGGAGAGAGAGAGAGCTTTCGAGAGGCTCTCTTC 660
QY 661 GAGCGGTGACCTCCGACGAGAAAGAAATACTCTGAGGAGGAGGAGGAGGAGGAGGAGG 720
DB 661 GAGCGGTGACCTCCGACGAGAAAGAAATACTCTGAGGAGGAGGAGGAGGAGGAGGAGG 720
QY 721 ATCAGCAGACGACGCTGCTTACCTCTGGGAGGCGGAGCTTAAACCGATACGAGAGG 780
DB 721 ATCAGCAGACGACGCTGCTTACCTCTGGGAGGCGGAGCTTAAACCGATACGAGAGG 780
QY 781 CCGCTGACGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
DB 781 CCGCTGACGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
QY 841 GCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 841 GCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 901 CACTGAGATTTACGGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
DB 901 CACTGAGATTTACGGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
```

```
QY 961 ACGAGATAGGCGAGCTCTACCTCTACGAGGCGGAGGTGAAGACAGGTACACCGAATACAC 1020
DB 961 ACGAGATAGGCGAGCTCTACCTCTACGAGGCGGAGGTGAAGACAGGTACACCGAATACAC 1020
QY 1021 GGGCCGATATTACGAGAGCTCAAGACCTTCGAGCCGAGGACCTTCGCTTCAAGACCAA 1080
DB 1021 GGGCCGATATTACGAGAGCTCAAGACCTTCGAGCCGAGGACCTTCGCTTCAAGACCAA 1080
QY 1081 GACCTGAGATTAAGGCTGCTACCTACGAGGCGGAGGTTAAAGAGAGAGGCGCGGAG 1140
DB 1081 GACCTGAGATTAAGGCTGCTACCTACGAGGCGGAGGTTAAAGAGAGAGGCGCGGAG 1140
QY 1141 ATAGCTTCTCCACGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
DB 1141 ATAGCTTCTCCACGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
QY 1201 CAGCTGATGCGAGCAAGGCGCTACTACTGCTGCTGTAACCCGCGGCGGAGCGAGC 1260
DB 1201 CAGCTGATGCGAGCAAGGCGCTACTACTGCTGCTGTAACCCGCGGCGGAGCGAGC 1260
QY 1261 TATAGCAAGACCTTCGCGCTCCGCTCCGCTGAGAGAGAGGAGGAGGAGGAGGAGGAG 1320
DB 1261 TATAGCAAGACCTTCGCGCTCCGCTCCGCTGAGAGAGAGGAGGAGGAGGAGGAGGAG 1320
QY 1321 ATATGAGACGAGATCGAGAGGATCTTCAAGCTGAAACCGAGGCGGAGGAGGAGGAGG 1380
DB 1321 ATATGAGACGAGATCGAGAGGATCTTCAAGCTGAAACCGAGGCGGAGGAGGAGGAGG 1380
QY 1381 GGAATTAACGGGATTAAGCTACGGGCGCTTATGACCAACATGGGCTTGAAGAGGAG 1440
DB 1381 GGAATTAACGGGATTAAGCTACGGGCGCTTATGACCAACATGGGCTTGAAGAGGAG 1440
QY 1441 CTCTTCAAGGAGGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
DB 1441 CTCTTCAAGGAGGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
QY 1501 TCGGACATAGGAGCTCTGTTACGAGCTGACGAGTCAATCGGCGCAATTCGTTAGAGAG 1560
DB 1501 TCGGACATAGGAGCTCTGTTACGAGCTGACGAGTCAATCGGCGCAATTCGTTAGAGAG 1560
QY 1561 AACTTCAGGAGGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
DB 1561 AACTTCAGGAGGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
QY 1621 ATCCACTGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
DB 1621 ATCCACTGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
QY 1681 CTCAAGGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
DB 1681 CTCAAGGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
QY 1741 AGCGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
DB 1741 AGCGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
QY 1801 GAGCGCAAGCTCAAGAGTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
DB 1801 GAGCGCAAGCTCAAGAGTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
QY 1861 GGGAGACTGA 1869
DB 1861 GGGAGACTGA 1869
```

RESULT 2
US-09-974-300-224
Sequence 224, Application US/09974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
APPLICANT: Clausen, Ib Groth
TITLE OF INVENTION: Methods for Monitoring Multiple Gene

RESULT 5
US-10-156-761-5055
: Sequence 5055, Application US/10156761
: Publication No. US20030119018A1
: GENERAL INFORMATION:
: APPLICANT: OMURA, SATOSHI
: APPLICANT: IKEDA, HARUO
: APPLICANT: ISHIKAWA, JUN
: APPLICANT: HORIKAWA, HIROSHI
: APPLICANT: SHIDA, TADAYOSHI
: APPLICANT: SAKAKI, YOSHIYUKI
: APPLICANT: HATTORI, MASAHIRA
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-262
: CURRENT APPLICATION NUMBER: US/10/156,761
: CURRENT FILING DATE: 2002-05-29
: PRIOR APPLICATION NUMBER: JP 2001-204089

```

RESULT 6
US-10-156-761-1
Sequence 1, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, YASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156, 761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 1
LENGTH: 9025608
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: misc_feature

```


QY	1077	AAAGACCTCGAGATAGACGGCTGTACCTTCAGGCCGGAAGGTTAAAGAGGAAGGCGCC	1136
Db	3447959	CGACATCACCGGCACCAATGTCATGAACACCTCTCTGGAAGTGTCTGCAGGCAAGCGCAC	34479000
QY	1137	GGTGTAGTCTTCCTCCACAGGCGGGCGGAGAGGGCATGTACGACACCGTTCCTACGA	1196
Db	3447899	CGTCGGCGCGACGTGTCTCGGACGGGCACTTCGGGGAAGCCTTCACCGCGTGGGGCTGT	3447840
QY	1197	GATCAGACTGATGGCGAGCAAGGGCTACTACTGCTCTTGCTGAACCCGGCGGACAGCA	1256
Db	3447839	CAACCCGCGGTGGCCCAACATGTCCTTCGCTTCGCGCGCGGGACGTGGCGCTCGGCGC	3447780
QY	1257	CGGCTATAGCAAGACATTCGCGCTCGGCGTCCG	1290
Db	3447779	CTTCTGGATCAGGGCGGGCGACGTCGTCTCCCG	3447746

```

1 RESULT 9
2 US-10-156-761-6210
3
4 Sequence 6210, Application US/10156761
5 Publication No. US20030119018A1
6 GENERAL INFORMATION:
7
8 APPLICANT: OMURA, SATOSHI
9 APPLICANT: IKEDA, HARUO
10 APPLICANT: ISHIKAWA, JUN
11 APPLICANT: HORIKAWA, HIROSHI
12 APPLICANT: SHIBA, TADAYOSHI
13 APPLICANT: SAKAKI, YOSHIYUKI
14 APPLICANT: HATTORI, MASAHIRA
15 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
16 FILE REFERENCE: 249-262
17
18 CURRENT APPLICATION NUMBER: US/10/156,761
19 PRIOR FILING DATE: 2002-05-29
20 PRIOR APPLICATION NUMBER: JP 2001-204089
21 PRIOR FILING DATE: 2001-05-30
22 PRIOR APPLICATION NUMBER: JP 2001-272697
23 PRIOR FILING DATE: 2001-08-02
24
25 NUMBER OF SEQ ID NOS: 15109
26
27 SEQ ID NO 6210
28
29 LENGTH: 1503
30
31 TYPE: DNA
32 ORGANISM: Streptomyces avermitilis
33 FEATURE:
34 NAME/KEY: CDS
35 LOCATION: (1)..(1503)
36 US-10-156-761-6210

```

Query Match	2.68;	Score 48.2;	DB 9;	Length 1503;
Best Local Similarity	43.78;	Pred. No. 0.00018;		
Matches 212;	Conservative 0;	Mismatches 273;	Indels 0;	Gaps 0;

QY	818	AGGCTCATCTTCCATCATCAATGCGGGCAGAGGTAACCTCTGGCTCTGGAGCGGGAAG	877
Db	65	AGGAGGCTTCTCTGCGCGCGATTCGACGAGACGATCACTACTTCAACAGCGGCATCTCG	124
QY	878	CCGAGGCTGTGTTACCGGGGACCACTGGATTTCAGGGCTTGAGCTCAGCATGGACAAG	937
Db	125	TCGACGGCGTCATCTGTAAAGGTCGACCGGGACGAGTCTCTGCATCGGTTCACAAAG	184
QY	938	CATGTCCTCATCATGACCGCCACAGAGATTAAGCGAGCTTACCTCTACGACGCGCGAG	997
Db	185	CCGAAAGGTGTATCTCCGAGCGCCGACGCTCTTGATCTACAGCAGATCTCAACCCGACGAGG	244
QY	998	TGAACAGATCAACGAAATACAAAGGGGCGGATATTTCAGGAAAGCTCAAGACTTTCGAGCGA	1057
Db	245	TCGTCAGAGTGGCGAGAGAGATTCGAAGCCCTTGTCTTCCAAAGAGGAGCAAGGAAGGACC	304
QY	1058	GGCACTTCCGCTTCAAGAGCAAAAGACCTTCGAGATAGAGGGGTGGTACTTCAGGCGGGAGG	1117
Db	305	GCCGTGATCTCTCGAAGAAAGCGGCCGCCACGTACAGAGCGGCTCTGGGGCCACCTTCGAGAAAGA	364
QY	1118	TTTAAAGAGAGAAAGGCCCGGCTGATTAGTCTTCTGCTCACGCGGGCCGAGAGGACATGTACG	1177

Db	365	TCAAGAGAGGACGGCATGTCACCGGACCGGTCATGAGGTGGTCAAGGGTGTC	424
Qy	1178	GAAACGGCTTCGTCTACGAGATGCAGCGATGAGGGAGAGAAAGGCTACTACTGGTCCTCG	123
Db	425	TCCTCAGACATCGGCTTCCTGGCTTCCTGCGGGCTTCCTCTGTGAAATGCGCGGGCTCC	484
Qy	1238	TGAACCCGGCGGCGACGACGAGCTTATACGGAAGCTTCGCGCTCCGGCTCTGGAGAGA	129
Db	485	GCGACCTCCAGGCTTCGTGGGCAAGGAGCTCGAGGCGAAGATCATGAGACTGAGACAAG	544
Qy	1298	CTGGC	1302
Db	545	ACCGC	549

```

RESULT 10
US-09-864-761-18040
Sequence 18040, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
FILE REFERENCE: Aecmlca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24463,6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 18040
LENGTH: 1785
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

```

Query Match	2.58;	Score 47;	DB 10;	Length 2428;
Best Local Similarity	49.48;	Pred. No. 0.00052;		
Matches 122; Conservative	0;	Mismatches 125;	Indels 0;	Gaps 0;

Query Match	2.5% Best Local Similarity	Score 47	DB 9	Length 2673
Matches 122	Conservative	0	Mismatches 125	Indels 0
			Gaps 0	
QY	1062	CTTCCGCTTCAGAGCAAGACCTCGAGATTAGACGGCTGTACTCCAGACCCGGAAGTTAA	1121	
Db	1419	CATGCCCCCTTGATATGGCCCTACGAAAGTCTCTTGGCCACAGACCTTACACCGTGGCCGT	1478	
QY	1122	AGAGGAGAAGGCCCCGGTGATAGCTTCCTGCTCCACGGGGGGCGGAAAGGCATGTACGACCA	1181	
Db	1479	CGGGCACTGTGTCTCTCTCATGTGCACGGCTCCGTGCTCTATGTCTTACCTGCTCTA	1538	
QY	1182	CCGCTTCTGTCATACGATGCAAGCTGATGGCGAGCAAGGGCTACTACTGCTCTTCGTGAA	1241	
Db	1539	TCTCTTCTTCGCGATGGCACAGCTGAGAGCAATTTCAAGGGCACCTTACGCTCTACTGTGCCC	1598	
QY	1242	CCGGCGGGGAGCAGCAGCGTATATACGAAAGATTCGCGCTCCGCGTCTGGAGAGATCGG	1301	
Db	1599	CTACCTGTGTGTGCTTCACTATGTGTGTAGCTCTCCGTGTCACTGTGTGGAGTCCACCGG	1658	
QY	1302	CTTGGAG	1308	
Db	1659	CTTGGGG	1665	

RESULT 3
US-10-156-761-2673
; Sequence 2673, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO

OY	1306	GAGACGCTTATGGACATAAAGAACCGCATCAGAGACTTTCTTAAGCTCGAACCCGAGGCC	1365
Db	42045	GCGGCGCTTCACGTCTGTACGAACTCGCTCGCGGCGAGTTACCAGACCCCTCGCTGGCGC	42104
OY	1366	GACACGAGAGCGCGGTGG	1382
Db	42105	GCGATTGCCGCGCGGCGG	42121

Search completed: July 15, 2003, 16:41:38
Job time : 334 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2003, 12:54:09 ; Search time 93 Seconds

(without alignments)
6163.214 Million cell updates/sec

Title: US-09-966-803-1

Sequence: 1 ATGACCGGATCGATGAGTGA.....TCAGGGGGAATGGAGACTGA 1869.

Scoring table: IDENTITY_NDC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1869	100.0	1869	2	US-08-664-646A-1
2	1869	100.0	1869	2	US-09-066-285-1
3	1869	100.0	1869	3	US-09-261-006-1
4	1869	100.0	1869	3	US-08-951-088-1
5	1869	100.0	1869	4	US-09-609-566-1
6	1869	100.0	1869	4	US-09-609-570-1
7	120.4	6.4	1896	3	US-09-016-080-2
8	48.2	2.6	1598	1	US-08-211-682-24
9	46.4	2.5	5977	3	US-09-024-020B-1
10	46.4	2.5	5977	3	US-09-425-043-1
11	46.4	2.5	6007	3	US-09-024-020B-2
12	46.4	2.5	6007	4	US-09-425-043-2
13	46.4	2.5	6556	3	US-09-024-020B-7
14	46.4	2.5	6556	4	US-09-425-043-7
15	46.4	2.5	6586	3	US-09-024-020B-43
16	46.4	2.5	6586	4	US-09-425-043-43
17	46.4	2.5	6826	3	US-09-024-020B-8
18	46.4	2.5	6826	4	US-09-425-043-8
19	46.2	2.5	7218	1	US-08-232-463-14
20	42.8	2.3	1299	2	US-08-761-277A-46
21	42.8	2.3	1299	4	US-09-199-637A-160
22	42.8	2.3	2051	4	US-08-338-219B-114
23	42.8	2.3	2241	3	US-08-838-219B-20
24	42.8	2.3	2241	3	US-09-233-336A-20
25	42.8	2.3	2241	3	US-09-233-752A-20
26	42.8	2.3	2241	4	US-09-402-036-20
27	42.8	2.3	2241	4	US-09-904-226-20

C	28	41.4	2.2	4403765	4	US-09-103-840A-2	Sequence 2, Appl
C	29	41.4	2.2	4411529	4	US-09-103-840A-1	Sequence 1, Appl
C	30	41.2	2.2	12588	2	US-08-387-942C-1	Sequence 1, Appl
C	31	40.6	2.2	3624	1	US-07-951-715A-6	Sequence 6, Appl
C	32	40.6	2.2	3624	2	US-08-459-448A-6	Sequence 6, Appl
C	33	40.6	2.2	3624	3	US-08-459-595A-6	Sequence 6, Appl
C	34	40.6	2.2	3624	3	US-08-459-504B-6	Sequence 6, Appl
C	35	40.6	2.2	3624	3	US-08-459-444-6	Sequence 6, Appl
C	36	40.6	2.2	3624	3	US-09-053-549-7	Sequence 7, Appl
C	37	40.6	2.2	3624	4	US-08-547-422-6	Sequence 6, Appl
C	38	40.2	2.2	1291	3	US-08-997-897-1	Sequence 1, Appl
C	39	40.2	2.2	1291	4	US-09-156-836B-1	Sequence 1, Appl
C	40	40.2	2.2	2852	4	US-09-056-556-203	Sequence 203, App
C	41	40.2	2.2	2852	4	US-08-843-417-1	Sequence 198, App
C	42	40.2	2.1	6344	4	US-08-659-656A-1	Sequence 1, Appl
C	43	40.2	2.1	6524	4	US-08-659-656A-7	Sequence 1, Appl
C	44	40.2	2.1	6527	4	US-08-659-656A-7	Sequence 7, Appl
C	45	40.2	2.1	7052	4	US-08-669-656A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-664-646A-1
Sequence 1, Application US/08664646A
Patent No. 5877001

GENERAL INFORMATION:

APPLICANT: Murphy et al.

TITLE OF INVENTION: Amides

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/664,646A

FILING DATE: June 17, 1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Charles J. Herron

REGISTRATION NUMBER: 28,019

REFERENCE/DOCKET NUMBER: 331400-53

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1744

TELEFAX: 201-994-1744

INFORMATION FOR SEQ. ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1869 NUCLEOTIDES

TYPE: NUCLEIC ACID

STRANDEDNESS: SINGLE

TOPOLOGY: LINEAR

MOLECULE TYPE: DNA

US-08-664-646A-1

Query Match 100.0%; Score 1869; DB 2; Length 1869;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGACCGGATCGATGAGTGAACGACGAGACCTTTTCTAAGTTCGCTACCTGGCGACCGG 60

Db 1 ATGACGCGCATCGAATGGAAACACAGAGACTTTTCTTAAGTTCCGCTTACTGGCGGACCCG 60
OY 61 AGGATACGGGGAAACTTAATCGCGTACACCTCGACGAAAGGCCAATGAAAGCAACAG 120
Db 61 AGGATACGGGGAAACTTAATCGCGTACACCTCGACGAAAGGCCAATGAAAGCAACAG 120
OY 121 TACGAGAGCAGCGTTGTTGTTGAAGACCTTGAACGGGGCTCAAGGGCTTCAATCGAGAC 180
Db 121 TACGAGAGCAGCGTTGTTGTTGAAGACCTTGAACGGGGCTCAAGGGCTTCAATCGAGAC 180
OY 181 GCGTCAATGCGGAGGATTTGCGCAGACGGGAGAAAGCTCGCTTCACTGCTTTAAGAG 240
Db 181 GCGTCAATGCGGAGGATTTGCGCAGACGGGAGAAAGCTCGCTTCACTGCTTTAAGAG 240
OY 241 GAGAGAGAGAGACCGAGATATGGGTGGCCGATATCCAGACCTGAGCGCCAAAGAAATGC 300
Db 241 GAGAGAGAGAGACCGAGATATGGGTGGCCGATATCCAGACCTGAGCGCCAAAGAAATGC 300
OY 301 CTCTCAACTTAAAAAGTCCGCTGATGATGAGAGACCGAGGTTTTCAGATGGGCTCTGAGCAG 360
Db 301 CTCTCAACTTAAAAAGTCCGCTGATGATGAGAGACCGAGGTTTTCAGATGGGCTCTGAGCAG 360
OY 361 GTGCGGCTTCAAGAGAGAGAGAGATGAGAGACTTGTCTTTGACGACGACGCTCCGCTG 420
Db 361 GTGCGGCTTCAAGAGAGAGAGAGATGAGAGACTTGTCTTTGACGACGACGCTCCGCTG 420
OY 421 TTGACAAATATGGGATTTCTTTGATGAGAGAAACGACGCTTCTGAGTCTGACACTGAG 480
Db 421 TTGACAAATATGGGATTTCTTTGATGAGAGAAACGACGCTTCTGAGTCTGAGTCTGAG 480
OY 481 GCCGAGAGATATATCGACAGCTTCGAGAACCGAGGTTTTCAGATGGGCTCTGAGCAGG 540
Db 481 GCCGAGAGATATATCGACAGCTTCGAGAACCGAGGTTTTCAGATGGGCTCTGAGCAGG 540
OY 541 GATGCGATATGTTGAAAGTCCCGCACCGGAGGGAGAGCAAGCTGCGCTTCAAGATTC 600
Db 541 GATGCGATATGTTGAAAGTCCCGCACCGGAGGGAGAGCAAGCTGCGCTTCAAGATTC 600
OY 541 GATGCGATATGTTGAAAGTCCCGCACCGGAGGGAGAGCAAGCTGCGCTTCAAGATTC 600
Db 541 GATGCGATATGTTGAAAGTCCCGCACCGGAGGGAGAGCAAGCTGCGCTTCAAGATTC 600
OY 601 TAGGACATAGTCTATGAGAGAGAGGGGAGAGAGAGGCTTTCGAGAGGGTCTCTCTTC 660
Db 601 TAGGACATAGTCTATGAGAGAGAGGGGAGAGAGAGGCTTTCGAGAGGGTCTCTCTTC 660
OY 661 GAGCGGATGACTCCGACGAGAAAGAGATCTCTGAGGGGCAAGAAAAAGCGGTTTC 720
Db 661 GAGCGGATGACTCCGACGAGAAAGAGATCTCTGAGGGGCAAGAAAAAGCGGTTTC 720
OY 721 ATCAGAGAGACGACACTGCTGCTGCTCTGAGAGGGGCAAGAAAAAGCGGTTTC 780
Db 721 ATCAGAGAGACGACACTGCTGCTGCTCTGAGAGGGGCAAGAAAAAGCGGTTTC 780
OY 781 CCGCTGACGCTGAGGAAAGCAAGCTCAGCGAAGAGAGGCTTACTCTCTCAGTCCAGAT 840
Db 781 CCGCTGACGCTGAGGAAAGCAAGCTCAGCGAAGAGAGGCTTACTCTCTCAGTCCAGAT 840
OY 841 GCGGGCAGAGGTTAAACCTCTGCTCTGAGGAGGAGGCGAGCGTGTGTTTACCGGAG 900
Db 841 GCGGGCAGAGGTTAAACCTCTGCTCTGAGGAGGAGGCGAGCGTGTGTTTACCGGAG 900
OY 901 CACTGATTTAAGCGGCTTGAAGCTGACGAGAGGAGGAGGATTTGCTCTCAGTACGAGCG 960
Db 901 CACTGATTTAAGCGGCTTGAAGCTGACGAGAGGAGGAGGATTTGCTCTCAGTACGAGCG 960
OY 961 ACAGAGATAGCAGAGCTCTACTCTTACGAGCGGAGCTGAAACAGTCAACGATCAAC 1020
Db 961 ACAGAGATAGCAGAGCTCTACTCTTACGAGCGGAGCTGAAACAGTCAACGATCAAC 1020
OY 1021 GGGCCGATATTCAGAGAGCTCAAGACCTTTCAGGCGAGGCACTTCGCTTCAAGAGCAA 1080
Db 1021 GGGCCGATATTCAGAGAGCTCAAGACCTTTCAGGCGAGGCACTTCGCTTCAAGAGCAA 1080
OY 1081 GAGCTGAGATTAAGAGGCTGCTACCTGAGGCGGAGGTTTAAAGAGAGAGGCGCGCGTG 1140
Db 1081 GAGCTGAGATTAAGAGGCTGCTACCTGAGGCGGAGGTTTAAAGAGAGAGGCGCGCGTG 1140

OY 1141 ATAGCTTCTGCTCCAGGCGGGGCGAAGGGCATGTACGAGACCGGCTTCTGTCAGAGATG 1200
Db 1141 ATAGCTTCTGCTCCAGGCGGGGCGAAGGGCATGTACGAGACCGGCTTCTGTCAGAGATG 1200
OY 1201 CAGCTATGGCGAGCAAGGGGCTACTAGTGTCTTGTGTAACCCCGGGCGGAGCGAGCG 1260
Db 1201 CAGCTATGGCGAGCAAGGGGCTACTAGTGTCTTGTGTAACCCCGGGCGGAGCGAGCG 1260
OY 1261 TATAGGAGAGACTTCCGCGCTCCGCTCTGAGAGAGAGAGTGTGAGAGACTTTGAGAG 1320
Db 1261 TATAGGAGAGACTTCCGCGCTCCGCTCTGAGAGAGAGAGTGTGAGAGACTTTGAGAG 1320
OY 1321 ATATGAGAGGAGCTCGAGAGGTTCTTCAAGCTCGAAGCCGAGCGCAGAGAGGCGGTT 1380
Db 1321 ATATGAGAGGAGCTCGAGAGGTTCTTCAAGCTCGAAGCCGAGCGCAGAGAGGCGGTT 1380
OY 1381 GGAATTAAGGGGCTAAGGCTACGCGGCTTCATGACCAAGCTGGGCTTGTACTGAGAGCG 1440
Db 1381 GGAATTAAGGGGCTAAGGCTACGCGGCTTCATGACCAAGCTGGGCTTGTACTGAGAGCG 1440
OY 1441 CTCTCAAGGAGAGAAATTAAGGAGAAAGGCTAAGCTACTGCTCAGAGCTACGCGTTTC 1500
Db 1441 CTCTCAAGGAGAGAAATTAAGGAGAAAGGCTAAGCTACTGCTCAGAGCTACGCGTTTC 1500
OY 1501 TCGGACATAGGCTGCTGATGAGAGAGGCTGAGAGGCTGAGGCTGAGGCTGAGGCTGAG 1560
Db 1501 TCGGACATAGGCTGCTGATGAGAGAGGCTGAGAGGCTGAGGCTGAGGCTGAGGCTGAG 1560
OY 1561 AACTTCAAGAGAGCTCAGAGGCTGCTGATGAGAGGCTGAGAGGCTGAGGCTGAGGCTGAG 1620
Db 1561 AACTTCAAGAGAGCTCAGAGGCTGCTGATGAGAGGCTGAGAGGCTGAGGCTGAGGCTGAG 1620
OY 1621 ATCCACTGCTTGAAGAGCTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Db 1621 ATCCACTGCTTGAAGAGCTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
OY 1681 CTCAGAGAGAGCTGAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1740
Db 1681 CTCAGAGAGAGCTGAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1740
OY 1741 AGCGTCCGCGAGAGCCCGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1800
Db 1741 AGCGTCCGCGAGAGCCCGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 1800
OY 1801 GAGCGAGAGCTCAAGAGAGTACGAGAGGCTTGAAGTGAAGAGTACTCAAGGGGAAT 1860
Db 1801 GAGCGAGAGCTCAAGAGAGTACGAGAGGCTTGAAGTGAAGAGTACTCAAGGGGAAT 1860
OY 1861 GGGAGCTGA 1869
Db 1861 GGGAGCTGA 1869

RESULT 2
US-09-285-1
; Sequence 1, Application us/09066285
; Patient No. 5985646
; GENERAL INFORMATION:
; APPLICANT: Murphy et al.
; TITLE OF INVENTION: Amides
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CABELLA, BYRNE, BAIN, GILLILAN,
; ADDRESS: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/066,285
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/664,646
 FILING DATE: June 17, 1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Charles J. Hepton
 REGISTRATION NUMBER: 28,019
 REFERENCE/DOCKET NUMBER: 331400-53
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1869 NUCLEOTIDES
 TYPE: NUCLEIC ACID
 STRANDEDNESS: SINGLE
 TOPOLOGY: LINEAR
 MOLECULE TYPE: DNA
 US-09-066-285-1

Query Match 100.0%; Score 1869; DB 2; Length 1869;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGGGCAATGGAACACAGAGACCTTTCTAAGTTCGCTACCTGGGCGACCG 60
 DB 1 ATGACGGGCAATGGAACACAGAGACCTTTCTAAGTTCGCTACCTGGGCGACCG 60
 QY 61 AGGATACGGGGAACCTTAACGGGTACACCTTACAGAAAGGCCAACAATGAACAAG 120
 DB 61 AGGATACGGGGAACCTTAACGGGTACACCTTACAGAAAGGCCAACAATGAACAAG 120
 QY 121 TACGAGACGACGGTGTGTGTTGAGACCTTGAAGAGGGCTCAAGGGGCTTCATCGAAG 180
 DB 121 TACGAGACGACGGTGTGTGTTGAGACCTTGAAGAGGGCTTCATCGAAG 180
 QY 181 GCCTCAATGCCAGATTTTCGCCAGACGGCAGAAAGCTGCTTCACTGCTTTAAACGAG 240
 DB 181 GCCTCAATGCCAGATTTTCGCCAGACGGCAGAAAGCTGCTTCACTGCTTTAAACGAG 240
 QY 241 GAGAAAGAGAGCCGAGATATGGGTGGCGGATATCCAGACCTGAGCCCAAGAAAGTC 300
 DB 241 GAGAAAGAGAGCCGAGATATGGGTGGCGGATATCCAGACCTGAGCCCAAGAAAGTC 300
 QY 301 CTCTCAATAAAAACGTCCTCGATGATGAGTGAAGAGATTCAGAGAGACTTAAGTT 360
 DB 301 CTCTCAATAAAAACGTCCTCGATGATGAGTGAAGAGATTCAGAGAGACTTAAGTT 360
 QY 361 GTCGGGTTCAAGAGAGAGAGATGAGACCTTCTTTGACGACGACGTCCTCCGGTCTGG 420
 DB 361 GTCGGGTTCAAGAGAGAGAGATGAGACCTTCTTTGACGACGACGTCCTCCGGTCTGG 420
 QY 421 TTCGACAATATGGGATCTTTGATGAGAGAGAGAGACGCTTCTGGGTTCTTACACTGAG 480
 DB 421 TTCGACAATATGGGATCTTTGATGAGAGAGAGAGACGCTTCTGGGTTCTTACACTGAG 480
 QY 481 GCCGAGAGATATGAGAGAGATTCGAGAGCCGAGGTTTCGAGTGGGCTTGGGACGGC 540
 DB 481 GCCGAGAGATATGAGAGAGATTCGAGAGCCGAGGTTTCGAGTGGGCTTGGGACGGC 540
 QY 541 GATCGGATAGTTGTGAAGCTCCCGACCGCGAGGGGAGACAACTGCTGTTCAAGTTC 600
 DB 541 GATCGGATAGTTGTGAAGCTCCCGACCGCGAGGGGAGACAACTGCTGTTCAAGTTC 600
 QY 601 TACGACATATGCTATGGAAGAGCGGGGAGAGAAAGCTTCTTGAAGAGGCTTCCTTC 660
 DB 601 TACGACATATGCTATGGAAGAGCGGGGAGAGAAAGCTTCTTGAAGAGGCTTCCTTC 660

QY 661 GAGCGGTTGATCTCCGACGGAAGAGATATCTCTGAGGGGCAAGAAAAAAGCGGTTTC 720
 DB 661 GAGCGGTTGATCTCCGACGGAAGAGATATCTCTGAGGGGCAAGAAAAAAGCGGTTTC 720
 QY 721 ATCAGGAGACGACGATGCTGCTGATCTTCTGGGAGGGGAGCTTAAACCGATCTACGAGGC 780
 DB 721 ATCAGGAGACGACGATGCTGCTGATCTTCTGGGAGGGGAGCTTAAACCGATCTACGAGGC 780
 QY 781 CCCTCGACGCTGAG 840
 DB 781 CCCTCGACGCTGAG 840
 QY 841 GCGGAG 900
 DB 841 GCGGAG 900
 QY 901 CACTGATTTAGGGGCTGAG 960
 DB 901 CACTGATTTAGGGGCTGAG 960
 QY 961 ACAGAGATAGGCGAGCTCTACCTCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 DB 961 ACAGAGATAGGCGAGCTCTACCTCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 QY 1021 GGGCCGATATTCAG 1080
 DB 1021 GGGCCGATATTCAG 1080
 QY 1081 GACCTCGAGATAG 1140
 DB 1081 GACCTCGAGATAG 1140
 QY 1141 ATAGCTTCTGTCACGAG 1200
 DB 1141 ATAGCTTCTGTCACGAG 1200
 QY 1201 CAGCTGATGCGGAG 1260
 DB 1201 CAGCTGATGCGGAG 1260
 QY 1261 TATAGGAGAGACTTCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1320
 DB 1261 TATAGGAGAGACTTCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1320
 QY 1321 ATATAG 1380
 DB 1321 ATATAG 1380
 QY 1381 GGAATTAAGGGGATTAAGCTACGCGGCTTCATGACCAACTGGGCTTGTACTAGAGCAG 1440
 DB 1381 GGAATTAAGGGGATTAAGCTACGCGGCTTCATGACCAACTGGGCTTGTACTAGAGCAG 1440
 QY 1441 CTCTTCAAGGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 DB 1441 CTCTTCAAGGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 QY 1501 TCGGACATATGGGCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
 DB 1501 TCGGACATATGGGCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
 QY 1561 AACTTCAG 1620
 DB 1561 AACTTCAG 1620
 QY 1621 ATCCACTGCTTGAAG 1680
 DB 1621 ATCCACTGCTTGAAG 1680
 QY 1681 CTCAAG 1740
 DB 1681 CTCAAG 1740
 QY 1741 AGCGTCCGCGAG 1800

1741 AGGTCGGGGAAGCCCGAGGCGCACAGGCCCAAGCGCTTACAGGCTCTTCAATAGAGTTCTTC 1800
QY 1801 GAGCGCAAGCTCAAGATGACGAGAGGCTTTAGGTAGAGAAATCTCAAGGGGAAT 1860
DB 1801 GAGCGCAAGCTCAAGATGACGAGAGGCTTTAGGTAGAGAAATCTCAAGGGGAAT 1860
QY 1861 GGGAACTGA 1869
DB 1861 GGGAACTGA 1869

RESULT 3
US-09-261-006-1
; Sequence 1, Application US/09261006
; Patent No. 6004796
; GENERAL INFORMATION:
; APPLICANT: Murphy et al.
; TITLE OF INVENTION: Amlases
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/261, 006
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/664, 646
; FILING DATE: June 17, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles J. Herion
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 331400-53
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ. ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1869 NUCLEOTIDES
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: DNA
; US-09-261-006-1

Query Match 100.0%; Score 1869; DB 3; Length 1869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AMGACGGGATGATGGAACGACGACCTTTTCTAAGTTGGCTTACCTGGGCGACCGG 60
DB 1 AMGACGGGATGATGGAACGACGACGACCTTTTCTAAGTTGGCTTACCTGGGCGACCGG 60
QY 61 AGGATACGGGGAAGCTTAATCGGCTACCTGACGAGGAGCAACATGGAAGGACACAAAG 120
DB 61 AGGATACGGGGAAGCTTAATCGGCTACCTGACGAGGAGCAACATGGAAGGACACAAAG 120
QY 121 TACGAGAGACGAGCTTGTGTTGAAGACCTTGAAGCGGCTCAAGGGCTTCAATCGAAGC 180
DB 121 TACGAGAGACGAGCTTGTGTTGAAGACCTTGAAGCGGCTCAAGGGCTTCAATCGAAGC 180
QY 1861 GCCTCAATGCGGAGATTTCGCCACAGCGCAGAAAGCTGCGCTTCACTGCTTTAAGAG 240

DB 181 GCCTCAATGCGGAGATTTCGCCACAGCGCAGAAAGCTGCGCTTCACTGCTTTAAGAG 240
QY 241 GAGAAAGAGAGACCGAGATATGGTGGCGATATCCAGACCCCTGAGCCCAAGAAATG 300
DB 241 GAGAAAGAGAGACCGAGATATGGTGGCGATATCCAGACCCCTGAGCCCAAGAAATG 300
QY 301 CTCTCAATCAAAAAGCTCCGCTCGATGCAAGTGAAGACGACGATTCAAGAGACTTTAGT 360
DB 301 CTCTCAATCAAAAAGCTCCGCTCGATGCAAGTGAAGACGACGATTCAAGAGACTTTAGT 360
QY 361 GTGCGCTTCAAGAGAGAGAGACGATGAGACTTCGTTTGAAGAGAGACGCTCCGCTG 420
DB 361 GTGCGCTTCAAGAGAGAGAGACGATGAGACTTCGTTTGAAGAGAGACGCTCCGCTG 420
QY 421 TTGCAATATGGGATTTCTTGAAGAGAGAGACGCTTCGAGTTCGAGACTGAG 480
DB 421 TTGCAATATGGGATTTCTTGAAGAGAGAGACGCTTCGAGTTCGAGACTGAG 480
QY 481 GCGGAGAGATATGAGAGAGATTCGAGAAAGCCGAGGTTTTCAGATGCTTGGCAGCGC 540
DB 481 GCGGAGAGATATGAGAGAGATTCGAGAAAGCCGAGGTTTTCAGATGCTTGGCAGCGC 540
QY 541 GATGCGATAGTTGTGAAGCTCCCGCAGCGAGGAGGAGCAAGCTGCTTCAAGTTC 600
DB 541 GATGCGATAGTTGTGAAGCTCCCGCAGCGAGGAGGAGCAAGCTGCTTCAAGTTC 600
QY 601 TACGACATAGCTTATGGAAGAGCGGGAGGAAGAGAACTTTCGAGAGGCTTCCTTC 660
DB 601 TACGACATAGCTTATGGAAGAGCGGGAGGAAGAGAACTTTCGAGAGGCTTCCTTC 660
QY 661 GAGGCGGTGATCCGAGGAGAAAGAAATACCTGAGGAGGAGAAAGGAGGAGGTTTC 720
DB 661 GAGGCGGTGATCCGAGGAGAAAGAAATACCTGAGGAGGAGAAAGGAGGAGGTTTC 720
QY 721 ATCAGGAGACGACTGCTGATCTCTGAGGAGCGGAGCTTAAACGATCTACAGAGGC 780
DB 721 ATCAGGAGACGACTGCTGATCTCTGAGGAGCGGAGCTTAAACGATCTACAGAGGC 780
QY 781 CCGCTGAGCTCGGAGGAGCAAGCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
DB 781 CCGCTGAGCTCGGAGGAGCAAGCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
QY 841 GCGGCGAGGATTAACCTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
DB 841 GCGGCGAGGATTAACCTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
QY 901 CACTGATTTACGCGCTGAGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
DB 901 CACTGATTTACGCGCTGAGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
QY 961 ACAGGATAGGAGAGCTACCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
DB 961 ACAGGATAGGAGAGCTACCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
QY 1021 GGGCGGATTTACGAGAGCTCAAGACCTTCAGGCGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1021 GGGCGGATTTACGAGAGCTCAAGACCTTCAGGCGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 GACCTGAGATAGAGCGGCTGATCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
DB 1081 GACCTGAGATAGAGCGGCTGATCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
QY 1141 ATAGCTTCTGCAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
DB 1141 ATAGCTTCTGCAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
QY 1201 CAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1201 CAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1261 TATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320

Db 1261 TATAGCAAGACTTCGGCTCCGGCTCTCGAGAGGACTGGCTTGAAGACTTTGAGGAC 1320
 QY 1321 ATATGAAAGCGCATCGAGAGTTCTTCAAGCTCGAACCGAGGCGGAGGCGGTT 1380
 Db 1321 ATATGAAAGCGCATCGAGAGTTCTTCAAGCTCGAACCGAGGCGGAGGCGGTT 1380
 QY 1381 GGAATATGAGGCGCATACCTACGAGGCGGCTTATGACCAACTGGGCTTGACTCAGAGGAC 1440
 Db 1381 GGAATATGAGGCGCATACCTACGAGGCGGCTTATGACCAACTGGGCTTGACTCAGAGGAC 1440
 QY 1441 CTCCTCAAGCAGGATATACGAGAGAGGCGATAGTACTGGCTCAGCAGCTTCCCTTC 1500
 Db 1441 CTCCTCAAGCAGGATATACGAGAGAGGCGATAGTACTGGCTCAGCAGCTTCCCTTC 1500
 QY 1501 TCGGACATAGGCGCTCTGTCAGACGTCGAGGTCATCGGGCCAAATCCGTTAGAGAACGAG 1560
 Db 1501 TCGGACATAGGCGCTCTGTCAGACGTCGAGGTCATCGGGCCAAATCCGTTAGAGAACGAG 1560
 QY 1561 AACTTCAGGAAGCTCAGCGCGCTGTCTAGGCTCAGAACGCTGAAGGCGCGATATCTTA 1620
 Db 1561 AACTTCAGGAAGCTCAGCGCGCTGTCTAGGCTCAGAACGCTGAAGGCGCGATATCTTA 1620
 QY 1621 ATCCACTCGCTTGAAGGACTACCGCTGTCCGCTGACACGAGGCTTATGTTCTCAACGTC 1680
 Db 1621 ATCCACTCGCTTGAAGGACTACCGCTGTCCGCTGACACGAGGCTTATGTTCTCAACGTC 1680
 QY 1681 CTCAGAGCATGTGGGCAAGGAGGCTTATAGCATATTCAGGCGGCGCCCAAGGCGCAC 1740
 Db 1681 CTCAGAGCATGTGGGCAAGGAGGCTTATAGCATATTCAGGCGGCGCCCAAGGCGCAC 1740
 QY 1741 AGCGTCGGGGAAGCCGAGGCGACAGGCGGACGCTTACAGGCTTCTTCAATAGGTTCTTC 1800
 Db 1741 AGCGTCGGGGAAGCCGAGGCGACAGGCGGACGCTTACAGGCTTCTTCAATAGGTTCTTC 1800
 QY 1801 GAGCGCAAGTCTCAAGAGTACGAGAGGCGCTTGAAGTACGAGGAGTCTCAAGGAGTAT 1860
 Db 1801 GAGCGCAAGTCTCAAGAGTACGAGAGGCGCTTGAAGTACGAGGAGTCTCAAGGAGTAT 1860
 QY 1861 GGGAACTGA 1869
 Db 1861 GGGAACTGA 1869

RESULT 4

US-08-951-088-1

Sequence 1, Application US/08951088

Patent No. 6136583

GENERAL INFORMATION:

APPLICANT: Murphy et al.

TITLE OF INVENTION: Amlidases

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESS: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/951,088

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/664,646

FILING DATE: June 17, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Charles J. Herion

REGISTRATION NUMBER: 28,019
 REFERENCE/DOCKET NUMBER: 331400-53
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1869 NUCLEOTIDES
 TYPE: NUCLEIC ACID
 STRANDEDNESS: SINGLE
 TOPOLOGY: LINEAR
 MOLECULE TYPE: DNA
 US-08-951-088-1

Query Match 100.0%; Score 1869; DB 3; Length 1869;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACCGGCATCGAATGGAACACAGAGACCTTTCTAAGTTCCCTTACTGAGGCGACCCG 60
 Db 1 ATGACCGGCATCGAATGGAACACAGAGACCTTTCTAAGTTCCCTTACTGAGGCGACCCG 60
 QY 61 AGGATACGGGGGAACCTTAATCGGGTACACCTTCAGGAAGGCCAATGAAAGCAACAG 120
 Db 61 AGGATACGGGGGAACCTTAATCGGGTACACCTTCAGGAAGGCCAATGAAAGCAACAG 120
 QY 121 TACGAGACGACGGTGTGTTGAAGACCTTGAACGGGCTCAAGGGGCTTCATCGAAGAC 180
 Db 121 TACGAGACGACGGTGTGTTGAAGACCTTGAACGGGCTCAAGGGGCTTCATCGAAGAC 180
 QY 181 GCCTCAATGCCAGAGATTTGCGCAGACGGCAGAAAGCTCCCTTCACTGCTTTAAGAG 240
 Db 181 GCCTCAATGCCAGAGATTTGCGCAGACGGCAGAAAGCTCCCTTCACTGCTTTAAGAG 240
 QY 241 GAGAAGAGAGACCGAGATATGTTGGGCGGATATTCAGACCTTGAGGCCAAGAAAGTC 300
 Db 241 GAGAAGAGAGACCGAGATATGTTGGGCGGATATTCAGACCTTGAGGCCAAGAAAGTC 300
 QY 301 CTCCTCACTAAACAGCTCCGCTCGATGATGACAGTGAAGAGATTCAGAGAGACTTAAATT 360
 Db 301 CTCCTCACTAAACAGCTCCGCTCGATGATGACAGTGAAGAGATTCAGAGAGACTTAAATT 360
 QY 361 GTGCGCTTCAAGAGGAGGACGATGAGAGCTTCTTGAAGAGAGAGCTCCCGGTCTGG 420
 Db 361 GTGCGCTTCAAGAGGAGGACGATGAGAGCTTCTTGAAGAGAGAGCTCCCGGTCTGG 420
 QY 421 TTCGACAAATATGGAGTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 Db 421 TTCGACAAATATGGAGTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 QY 481 GCCGAGAGATATATCGAGAGTTCGAGAAAGCCGAGGTTTTCGAGTGGCTTGGACAGGCG 540
 Db 481 GCCGAGAGATATATCGAGAGTTCGAGAAAGCCGAGGTTTTCGAGTGGCTTGGACAGGCG 540
 QY 541 GATGCGATATGTTGTAAGTCCCGACCGCGAGGGAGAGCAAGCTTCCCTGTTCAAGTTTC 600
 Db 541 GATGCGATATGTTGTAAGTCCCGACCGCGAGGGAGAGCAAGCTTCCCTGTTCAAGTTTC 600
 QY 601 TACGACATATGCTTATGGAAGACGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 Db 601 TACGACATATGCTTATGGAAGACGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 QY 661 GAGCGGGTGTGACTCCGACGGAAGAAATATCTCTGAGGGGCAAGAAAAAAGCGGTTTC 720
 Db 661 GAGCGGGTGTGACTCCGACGGAAGAAATATCTCTGAGGGGCAAGAAAAAAGCGGTTTC 720
 QY 721 ATCAGCGAGACAGACTGCTGTGTAAGTCTGAGAGCGGCGAGCTTAAACCGATTCAGAGGCG 780
 Db 721 ATCAGCGAGACAGACTGCTGTGTAAGTCTGAGAGCGGCGAGCTTAAACCGATTCAGAGGCG 780
 QY 781 CCGCTCGACGCTGGGGAAGCCAAAGCTCACGGAGGAAGAGTCTACTCTCTCACTCAAGAT 840
 Db 781 CCGCTCGACGCTGGGGAAGCCAAAGCTCACGGAGGAAGAGTCTACTCTCTCACTCAAGAT 840

QY 841 GGGGGAGGGTAACCTCTGGCTCTGGAGGGGAGCCGAGCGTCTTGTACCGCGAC 900
|||||
DB 841 GGGGGAGGGTAACCTCTGGCTCTGGAGGGGAGCCGAGCGTCTTGTACCGCGAC 900
QY 901 CACTGATTTAGGGGCTTACGCTCAGCGATGGCAAGCAATGCTCTCATCATAGACGGC 960
|||||
DB 901 CACTGATTTAGGGGCTTACGCTCAGCGATGGCAAGCAATGCTCTCATCATAGACGGC 960
QY 961 ACAGAGATAGGGAGCTCTACCTCTACGAGCGGAGCTTAACAGTCAACGAATPACAC 1020
|||||
DB 961 ACAGAGATAGGGAGCTCTACCTCTACGAGCGGAGCTTAACAGTCAACGAATPACAC 1020
QY 1021 GGGCGGATTTAGAGAAAGCTCAAGACCTTCAGCCGAGCAGCCTTCGCTTCAAGACAAA 1080
|||||
DB 1021 GGGCGGATTTAGAGAAAGCTCAAGACCTTCAGCCGAGCAGCCTTCGCTTCAAGACAAA 1080
QY 1081 GACCTCGAGATAGAGCGCTGTACCTCAGAGCCGAGGCTTAAGAGAGAGAGCCCGGCTG 1140
|||||
DB 1081 GACCTCGAGATAGAGCGCTGTACCTCAGAGCCGAGGCTTAAGAGAGAGAGCCCGGCTG 1140
QY 1141 ATAGCTTCCTCCAGCGCGCGCGGCGAAGGCGATGACGACACCGCTTCCTACGAGATG 1200
|||||
DB 1141 ATAGCTTCCTCCAGCGCGCGCGGCGAAGGCGATGACGACACCGCTTCCTACGAGATG 1200
QY 1201 CAGCTGATGCGAGGAGGAGGCTACTACTGCTGCTGTAACCCGCGCGAGCGAGCGG 1260
|||||
DB 1201 CAGCTGATGCGAGGAGGAGGCTACTACTGCTGCTGTAACCCGCGCGAGCGAGCGG 1260
QY 1261 TATAGGAGAGACTTCGCGCTCCGCTCCTGAGAGAGAGCTGGCTTGAGAGACTTTGAGAG 1320
|||||
DB 1261 TATAGGAGAGACTTCGCGCTCCGCTCCTGAGAGAGAGCTGGCTTGAGAGACTTTGAGAG 1320
QY 1321 ATAAAGAGAGGAGATGAGAGATTTCTCAAGCTCGAAGCCGAGCGCAGAGGAGCGCTT 1380
|||||
DB 1321 ATAAAGAGAGGAGATGAGAGATTTCTCAAGCTCGAAGCCGAGCGCAGAGGAGCGCTT 1380
QY 1381 GGAATPACGGGAGATAGCTACGCGGCTTCATACCACTGGGCGCTTGACTGAGAGCGAC 1440
|||||
DB 1381 GGAATPACGGGAGATAGCTACGCGGCTTCATACCACTGGGCGCTTGACTGAGAGCGAC 1440
QY 1441 CTCTTCAGAGGAGAGATTAAGCAGAGAGGCGATTAAGCTACTGGCTCACAGCTACGCGCTTC 1500
|||||
DB 1441 CTCTTCAGAGGAGAGATTAAGCAGAGAGGCGATTAAGCTACTGGCTCACAGCTACGCGCTTC 1500
QY 1501 TCGGACATAGGGGCTCTGGTACGAGCGCTGAGGCTCATGCGGCAAAATCCGTTAGAGAGAG 1560
|||||
DB 1501 TCGGACATAGGGGCTCTGGTACGAGCGCTGAGGCTCATGCGGCAAAATCCGTTAGAGAGAG 1560
QY 1561 AACTTCAGAGAGCTCAGCCGCTGTCTACGCTCAGAGAGGCTGAAGGCGCGAGATCTCTA 1620
|||||
DB 1561 AACTTCAGAGAGCTCAGCCGCTGTCTACGCTCAGAGAGGCTGAAGGCGCGAGATCTCTA 1620
QY 1621 ATCCACTCGCTTGAAGAGTACCGCTGTCCGCTGACAGAGCGCTTATGTTCTACAGCGTG 1680
|||||
DB 1621 ATCCACTCGCTTGAAGAGTACCGCTGTCCGCTGACAGAGCGCTTATGTTCTACAGCGTG 1680
QY 1681 CTCAGAGAGATGGGAG 1740
|||||
DB 1681 CTCAGAGAGATGGGAG 1740
QY 1741 AGCGTCCGCGAG 1800
|||||
DB 1741 AGCGTCCGCGAG 1800
QY 1801 GAGCGGAGAGCTCAAGAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
|||||
DB 1801 GAGCGGAGAGCTCAAGAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
QY 1861 GGGAGACTGA 1869
|||||
DB 1861 GGGAGACTGA 1869

RESULT 5
US-09-609-566-1
; Sequence 1, Application us/09609566
; Patent No. 6429004
GENERAL INFORMATION:
; APPLICANT: Murphy et al.
; TITLE OF INVENTION: Amidases
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/609,566
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 09/427,372
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles J. Heiron
; REGISTRATION NUMBER: 28,019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1869 NUCLEOTIDES
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: DNA
US-09-609-566-1
Query Match 100.0%; Score 1869; DB 4; Length 1869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACCGGAGATGAGATGAG 60
|||||
DB 1 ATGACCGGAGATGAGATGAG 60
QY 61 AGGATACGGGGAGAACTTAATCGCTACACCCCTGACGAGAGGCCAGATGAGAGACAAAG 120
|||||
DB 61 AGGATACGGGGAGAACTTAATCGCTACACCCCTGACGAGAGGCCAGATGAGAGACAAAG 120
QY 121 TACGAGAGACAGCGTGTGTTTAAGACCTTTGAACCGGCTCAAGCGCTTACTACGAGAAC 180
|||||
DB 121 TACGAGAGACAGCGTGTGTTTAAGACCTTTGAACCGGCTCAAGCGCTTACTACGAGAAC 180
QY 181 GCCCTAATCCGAGAGATTTGCGCAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
|||||
DB 181 GCCCTAATCCGAGAGATTTGCGCAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 GAGAGAGAGAGAGAGAGATATGAGTGGCCGATATCCAGACCTTGAAGCGCCAGAGAAATC 300
|||||
DB 241 GAGAGAGAGAGAGAGAGATATGAGTGGCCGATATCCAGACCTTGAAGCGCCAGAGAAATC 300
QY 301 CTCTCACTAAAGAGTCCGCTCGATGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
|||||
DB 301 CTCTCACTAAAGAGTCCGCTCGATGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY	361	GTGGGCTTCAAGAGGAGGAGCAGATGAGACCTGTGGCTTTACACAGACGCTCCGGTCTGG	420
Db	361	GTGGGCTTCAAGAGGAGGAGCAGATGAGACCTGTGGCTTTACACAGACGCTCCGGTCTGG	420
QY	421	TTGCACAAATATGAGGATTCCTTTGATGAGAGAGACGACGCTTCTGGGCTTTGACACTGAG	480
Db	421	TTGCACAAATATGAGGATTCCTTTGATGAGAGAGACGACGCTTCTGGGCTTTGACACTGAG	480
QY	481	GGCAGAGAGATATATCAGACGAGTTCGAGAAACCCAGAGTTTTCGAGTGGGCTCTGGCAGGC	540
Db	481	GGCAGAGAGATATATCAGACGAGTTCGAGAAACCCAGAGTTTTCGAGTGGGCTCTGGCAGGC	540
QY	541	GATGCGATAGTTGTGAACGTCGCCGACCCGGAGAGGGGAGCAAGCGTCGCCCTTCAAGTTTC	600
Db	541	GATGCGATAGTTGTGAACGTCGCCGACCCGGAGAGGGGAGCAAGCGTCGCCCTTCAAGTTTC	600
QY	601	TACGACATATGTCCTATATGAGAGAGCAGGGGAGAGAAAGAAAGTCTTTCGAGAGGGTCTCTTC	660
Db	601	TACGACATATGTCCTATATGAGAGAGCAGGGGAGAGAAAGAAAGTCTTTCGAGAGGGTCTCTTC	660
QY	661	GAGGCGGTTGACTCCGACGGAAAGAGAAATCTCCTGAGGGGCGAAGAAAAAGGGGTTTC	720
Db	661	GAGGCGGTTGACTCCGACGGAAAGAGAAATCTCCTGAGGGGCGAAGAAAAAGGGGTTTC	720
QY	721	ATCAGCGAGACGAGCTGGGCTGTACCTCTGGGAGCGGAGCTTAAACGATCTACAGAGGC	780
Db	721	ATCAGCGAGACGAGCTGGGCTGTACCTCTGGGAGCGGAGCTTAAACGATCTACAGAGGC	780
QY	781	CCGCTTGAGCTCGGGAGACCAAGCTCACGGAAAGAAAGTCTACTCTCATCTCCATCCAGAT	840
Db	781	CCGCTTGAGCTCGGGAGACCAAGCTCACGGAAAGAAAGTCTACTCTCATCTCCATCCAGAT	840
QY	841	GGGGGACGGGTAAACCTCTGGGCTCTGGGAGCGGAGAACCCGAGGCTTTGTTCACCGCGAC	900
Db	841	GGGGGACGGGTAAACCTCTGGGCTCTGGGAGCGGAGAACCCGAGGCTTTGTTCACCGCGAC	900
QY	901	CAGTGGATTTAGGGGCTTGACGTCACCGAATGGCAAGCATGTGCTCTCATCATGACCGCC	960
Db	901	CAGTGGATTTAGGGGCTTGACGTCACCGAATGGCAAGCATGTGCTCTCATCATGACCGCC	960
QY	961	ACGAGGATGAGGAGGACTTACCTCTACGAGCGGAGAGCTAAACAGGTCACCGAATATCAAC	1020
Db	961	ACGAGGATGAGGAGGACTTACCTCTACGAGCGGAGAGCTAAACAGGTCACCGAATATCAAC	1020
QY	1021	GGGGCGATTTTGAAGAAAGCTCAAGACCTTCGAGCCGAGGACTTCCGCTTCAAGACCAA	1080
Db	1021	GGGGCGATTTTGAAGAAAGCTCAAGACCTTCGAGCCGAGGACTTCCGCTTCAAGACCAA	1080
QY	1081	GACCTCGAGATGAGCGGCTGATACCTCAGCCGAGAGGTTAAAGAGAGAAAGGCCCGCGGTG	1140
Db	1081	GACCTCGAGATGAGCGGCTGATACCTCAGCCGAGAGGTTAAAGAGAGAAAGGCCCGCGGTG	1140
QY	1141	ATATGCTTCTGTCCACGGGCGGCGCGAAGGGCATGTACGACACCGCTTCGTTACAGAGATG	1200
Db	1141	ATATGCTTCTGTCCACGGGCGGCGCGAAGGGCATGTACGACACCGCTTCGTTACAGAGATG	1200
QY	1201	CAGCTGATGGCGAGCAAGGGGCTACATCTGTCTCGTGTAAACCCCGCGGGAGCGCAAGCGC	1260
Db	1201	CAGCTGATGGCGAGCAAGGGGCTACATCTGTCTCGTGTAAACCCCGCGGGAGCGCAAGCGC	1260
QY	1261	TATATCGAAGACTTCCGCTCCGCTCCGCTCGAGAGAGACTGGCTTGAAGAGACTTTGAGAGC	1320
Db	1261	TATATCGAAGACTTCCGCTCCGCTCCGCTCGAGAGAGACTGGCTTGAAGAGACTTTGAGAGC	1320
QY	1321	ATATAGAACGGCATGAGAGAGTTCCTTCAAGCTCGAACCAGGAGCCGACAGAGAGGCGGCTT	1380
Db	1321	ATATAGAACGGCATGAGAGAGTTCCTTCAAGCTCGAACCAGGAGCCGACAGAGAGGCGGCTT	1380
QY	1381	GGAAATACGGGCGATATAGCTACGGCGGCTTCAATGACCACTGGGCTTACTCATGAGCGAC	1440
Db	1381	GGAAATACGGGCGATATAGCTACGGCGGCTTCAATGACCACTGGGCTTACTCATGAGCGAC	1440
QY	1441	CTCTTCAAGGCAAGGAATAGCAGAGACGGCATATAGCTACTGGCTTACCAAGCTACGCTTTC	1500

Db 1441 CTCTTCAAGGAGGAAATTAGGAGAAAGGCGCTAAAGTACTGCTCACCAGCTACGCCCTTC 15000
 QY 1501 TCGGACATAGGGCTCTGTGTACGAGCTGAGGTCTATCGGGCCCAAAATCCGTTTAGAAGAGAG 15600
 Db 1501 TCGGACATAGGGCTCTGTGTGTACGAGCTGAGGTCTATCGGGCCCAAAATCCGTTTAGAAGAGAG 15600
 QY 1561 AACTTTCAGGAAGCTTCAGGCCGCGCTGTCTTACGCTCAGAAAGCTGAAGGCGCGATACCTCTA 16200
 Db 1561 AACTTTCAGGAAGCTTCAGGCCGCGCTGTCTTACGCTCAGAAAGCTGAAGGCGCGATACCTCTA 16200
 QY 1621 ATTCACATCGCTTGAGAGACTACCGCTGTCCGCTGCACCAAGAGACCTTAATGCTTACAAAGTG 16800
 Db 1621 ATTCACATCGCTTGAGAGACTACCGCTGTCCGCTGCACCAAGAGACCTTAATGCTTACAAAGTG 16800
 QY 1681 CTCAAGGACATATGGGCAAGAGAAAGCTTACATATAGCATATTTCAAGCGCGGCGCCACAGGCGCAC 17400
 Db 1681 CTCAAGGACATATGGGCAAGAGAAAGCTTACATATAGCATATTTCAAGCGCGGCGCCACAGGCGCAC 17400
 QY 1741 AGCGTCGCGGAAAGCCCGAGGCAACAGGCGCTACAGGCTCTTCATATAGATTTCTTC 18000
 Db 1741 AGCGTCGCGGAAAGCCCGAGGCAACAGGCGCTACAGGCTCTTCATATAGATTTCTTC 18000
 QY 1801 GAGCGCAAGCTCAAGAAAGTACGAGAGAGGCGCTTTGAGGTAGAGAAAGATATCTCAAGGGGAAT 18600
 Db 1801 GAGCGCAAGCTCAAGAAAGTACGAGAGAGGCGCTTTGAGGTAGAGAAAGATATCTCAAGGGGAAT 18600
 QY 1861 GGGAACTGA 1869
 Db 1861 GGGAACTGA 1869

RESULT 6
 US-09-609-570-1
 Sequence 1, Application US/09609570
 Patent No. 6465204
 GENERAL INFORMATION:
 APPLICANT: Murphy et al.
 TITLE OF INVENTION: Amldases
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN
 CECCHI, STEWART & OLSTEIN
 STREET: 6 BECKER FARM ROAD
 CITY: ROSELAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/609,570
 FILING DATE: 30-Jun-2000
 CLASSIFICATION: <unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/427,372
 FILING DATE: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Charles J. Herion
 REGISTRATION NUMBER: 28,019
 REFERENCE/DOCKET NUMBER: 331400-53
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1869 NUCLEOTIDES
 TYPE: NUCLEIC ACID
 STRANDEDNESS: SINGLE
 TOPOLOGY: LINEAR

```

;      MOLECULE TYPE: DNA
;      SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-609-570-1

```

Query Match	100.0%;	Score 1869;	DB 4;	Length 1869;
Best Local Similarity	100.0%;	Pred. NO. 0;		
Matches 1869;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGACGGCATCGAATGGACACAGACACTTTTCTAGTTGGCTGCTACTGGGCGACCG	60
Db	1	ATGACCGGCATCGAATGGACACAGACACTTTTCTAGTTGGCTGCTACTGGGCGACCG	60
QY	61	AGGATACGGGAAACTTAATCGCGTACACCCCTGACGAAGGCCCAATGAAAGCAACAAG	120
Db	61	AGGATACGGGAAACTTAATCGCGTACACCCCTGACGAAGGCCCAATGAAAGCAACAAG	120
QY	121	TACGAGAGCACGCGTTGTTGTTGAAGACCTTGAAACGGGCTCAAGGCGCTTCATCGAGAC	180
Db	121	TACGAGAGCACGCGTTGTTGTTGAAGACCTTGAAACGGGCTCAAGGCGCTTCATCGAGAC	180
QY	181	GCCCATATGCCGAGAGATTGTGCGACAGCGGCAAGAAAGCTCGGCTTCACCTGCTTTAAGAG	240
Db	181	GCCCATATGCCGAGAGATTGTGCGACAGCGGCAAGAAAGCTCGGCTTCACCTGCTTTAAGAG	240
QY	241	GAGAAGAAGAGACCGAGATATGGTGGCCGATATTCAGACCTGAGGCCCAAGAAATTC	300
Db	241	GAGAAGAAGAGACCGAGATATGGTGGCCGATATTCAGACCTGAGGCCCAAGAAATTC	300
QY	301	CTCTCAACTTAAAAACGTCCGCTCGATGCAATGAGAACGACGATTCAGAGAGACTCTTAATT	360
Db	301	CTCTCAACTTAAAAACGTCCGCTCGATGCAATGAGAACGACGATTCAGAGAGACTCTTAATT	360
QY	361	GTGCGCTTCAAGAGAGAGGAGATGAGGACTTCGCTTTGACAGACAGTCCCGGTCTGG	420
Db	361	GTGCGCTTCAAGAGAGAGGAGATGAGGACTTCGCTTTGACAGACAGTCCCGGTCTGG	420
QY	421	TTGCACAATATGGAATCTTGTATGAGAGAGAAAGACGACGTTCTGGGTTCTTGACACTGAG	480
Db	421	TTGCACAATATGGAATCTTGTATGAGAGAGAAAGACGACGTTCTGGGTTCTTGACACTGAG	480
QY	481	GCGGAGAGATAAATCGAGACATTCGAGAAAGCCGAGGTTTGTGAGTGGGCTCTGGACGGC	540
Db	481	GCGGAGAGATAAATCGAGACATTCGAGAAAGCCGAGGTTTGTGAGTGGGCTCTGGACGGC	540
QY	541	GATGCGATAGTGTGAACGTCGCCGACCGCAGGAGGAGCAAGACCTGCGCTTCAAGTTTC	600
Db	541	GATGCGATAGTGTGAACGTCGCCGACCGCAGGAGGAGCAAGACCTGCGCTTCAAGTTTC	600
QY	601	TACGACATAGTCTATGAGAGAGCGGGAGGAAGAAAGCTCTTTCGAGAGGGTCTCTTTC	660
Db	601	TACGACATAGTCTATGAGAGAGCGGGAGGAAGAAAGCTCTTTCGAGAGGGTCTCTTTC	660
QY	661	GAGCGGGTGACTCCGACGAGAAAGAAATCTCTGAGGGGCAAGAAAAAAGGCTTC	720
Db	661	GAGCGGGTGACTCCGACGAGAAAGAAATCTCTGAGGGGCAAGAAAAAAGGCTTC	720
QY	721	ATCAGCAGACGACGACTGGCTGTACTCTGGGACGGCGAGACTTAAACGATATACGAGGC	780
Db	721	ATCAGCAGACGACGACTGGCTGTACTCTGGGACGGCGAGACTTAAACGATATACGAGGC	780
QY	781	CCGCTCGACGCTCTGGGAAAGCCAGCTCACGGAAGGAAAGTCTACTCTCACTCAGAT	840
Db	781	CCGCTCGACGCTCTGGGAAAGCCAGCTCACGGAAGGAAAGTCTACTCTCACTCAGAT	840
QY	841	GCGGCGAGGGTAAACCTCTGGCTCTGGGAGCGGGAAGGCGGAGCGTGTATACGGGCGAC	900
Db	841	GCGGCGAGGGTAAACCTCTGGCTCTGGGAGCGGGAAGGCGGAGCGTGTATACGGGCGAC	900
QY	901	CACGTGAATTTACGGGCTTGACGTACGCGATGGCAAGCATTTGCTCTCATCATGACGCGC	960
Db	901	CACGTGAATTTACGGGCTTGACGTACGCGATGGCAAGCATTTGCTCTCATCATGACGCGC	960
QY	961	ACGAGGATAGCGAGCTCTACTCTTACGAGCGCGAGCTGAACAGGTCACCGAATACAC	1020

Db	961	ACGAGGATATGACCGAGCTCTACCTCTACGACGCGCAGCTGAAACAGGTCACCGAATATACAC	1020
Qy	1021	GGGCGGATATTCAGAGAGCTCAACACCTTCGAGCGCGAGGCATTCGCGCTTCAGAGCAA	1080
Db	1021	GGGCGGATATTCAGAGAGCTCAACACCTTCGAGCGCGAGGCATTCGCGCTTCAGAGCAA	1080
Qy	1081	GACCTCGAGATAGACGGCTGTGTACTCAGCGCGGAGGTTAAAGAGAAAGGCCCGCGTG	1140
Db	1081	GACCTCGAGATAGACGGCTGTGTACTCAGCGCGGAGGTTAAAGAGAAAGGCCCGCGTG	1140
Qy	1141	ATAGTCTTCGTCCACGGCGGGCCGAAAGGCAATGACGGACACCGCTTCGTCTACGAGATG	1200
Db	1141	ATAGTCTTCGTCCACGGCGGGCCGAAAGGCAATGACGGACACCGCTTCGTCTACGAGATG	1200
Qy	1201	CAGCTGATGGGAGACAGGGGTACTACTGCTGCTTCTGTGAACCCGCGCGGACAGGACGGC	1260
Db	1201	CAGCTGATGGGAGACAGGGGTACTACTGCTGCTTCTGTGAACCCGCGCGGACAGGACGGC	1260
Qy	1261	TATAGCAAGACTTCGGGCTCCGGCTCCTGGAGAGACTGGCTTGGAGAGATTTGAGAGAC	1320
Db	1261	TATAGCAAGACTTCGGGCTCCGGCTCCTGGAGAGACTGGCTTGGAGAGATTTGAGAGAC	1320
Qy	1321	ATAATGAAACGGCATCGAGAGATTCTTCAAGCTCGAACCCGACGCGCGACAGGGACGGCTT	1380
Db	1321	ATAATGAAACGGCATCGAGAGATTCTTCAAGCTCGAACCCGACGCGCGACAGGGACGGCTT	1380
Qy	1381	GGAAATACGGGCAATTAACCTACGGGGGCTTCAATGACCAACATGGGCTTGACTCAAGAGGAC	1440
Db	1381	GGAAATACGGGCAATTAACCTACGGGGGCTTCAATGACCAACATGGGCTTGACTCAAGAGGAC	1440
Qy	1441	CTCTTCAAGCGAGGAATTAAGCGAAGAACGGCATAAGCTACTGGCTCAACAGCTACGCCCTTC	1500
Db	1441	CTCTTCAAGCGAGGAATTAAGCGAAGAACGGCATAAGCTACTGGCTCAACAGCTACGCCCTTC	1500
Qy	1501	TCGCAKATAGGGCTCTGCTGTACAGCTGAGGTCAATCGGGCCAAATCCGCTTAGAGAACGAG	1560
Db	1501	TCGCAKATAGGGCTCTGCTGTACAGCTGAGGTCAATCGGGCCAAATCCGCTTAGAGAACGAG	1560
Qy	1561	AACCTTCAGGAAGCTCAGCGCGCTTTCAGCGCTCAGAACGTAAGGAGGCGCATACTCTTA	1620
Db	1561	AACCTTCAGGAAGCTCAGCGCGCTTTCAGCGCTCAGAACGTAAGGAGGCGCATACTCTTA	1620
Qy	1621	ATCCACTCGCTTGAGGACTACCGCTGTCCGCTCGACCAAGACCTTATGTTCTACAAAGTG	1680
Db	1621	ATCCACTCGCTTGAGGACTACCGCTGTCCGCTCGACCAAGACCTTATGTTCTACAAAGTG	1680
Qy	1681	CTCAAGACATGGGCAAGAAAGCTTAATATAGCATATTCAAAGCGCGCGCCACAGGCGCAC	1740
Db	1681	CTCAAGACATGGGCAAGAAAGCTTAATATAGCATATTCAAAGCGCGCGCCACAGGCGCAC	1740
Qy	1741	AGCGTCCGCGGAACCCGAGGACACAGGCGCGAAGGCTTACAGGCGCTTCATAGATTTCTTC	1800
Db	1741	AGCGTCCGCGGAACCCGAGGACACAGGCGCGAAGGCTTACAGGCGCTTCATAGATTTCTTC	1800
Qy	1801	GAGCGCAAGCTCAAGAAAGTACGAGAGGCGCTTGAAGTAGAGAAAGATACCTCAAGGGGAAT	1860
Db	1801	GAGCGCAAGCTCAAGAAAGTACGAGAGGCGCTTGAAGTAGAGAAAGATACCTCAAGGGGAAT	1860
Qy	1861	GGGAACCTGA 1869	
Db	1861	GGGAACCTGA 1869	

RESULT 7
US-09-016-080-2
Sequence 2, Application US/090160800
Patent No. 6133012
GENERAL INFORMATION:
APPLICANT: Ishikawa, Kazuhiko
APPLICANT: Matsui, Ikuo
APPLICANT: Ishida, Hiroyasu
APPLICANT: Kosugi, Yoshitsugu

APPLICANT: HERMAN, RONALD C.
APPLICANT: SANGAMESWARAN, LAKSHMI
TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: JANET PAULINE CLARK
STREET: 3401 HILLYER AVENUE, MS A2-250
CITY: PALO ALTO
STATE: CA
COUNTRY: U.S.A.
ZIP: 94304-1397
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,020B
FILING DATE: 16-FEB-1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,447
FILING DATE: 26-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: CLARK, JANET P.
REGISTRATION NUMBER: 34,799
REFERENCE/DOCKET NUMBER: R0020B-REG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 852-3097
TELEFAX: (650) 855-5322
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6007 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-024-020B-2

Query Match 2.5%; Score 46.4; DB 3; Length 6007;
Best Local Similarity 48.2%; Pred. No. 0.029;
Matches 131; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

DB 1200 GCAGCTGATGGCGAGCAAGGCGTACTGCTGCTGTAACCCGCGGCGAGCGG 1259
4176 GAAGCTCATGAGGAGGCAACAGCAGAGATCCGATGGAGAAGATGTAAGATCAACTTTGA 4235

QY 1260 CTATAGCGAAGACTGCGCGCTCGGCTGGAGAGAGACTGGCTTGAGAGACTTTGAGGA 1319
DB 4236 CATATGTCGAGGAGGAGGACTGCGCTTCTTCAAGTGGCAACCTTCAAAAGCTGGATGGA 4295

QY 1320 CATATGAGACGGCATGAGAGAGTTCTTCAAGCTGGAACCGCAGCGGAGAGGCGCGT 1379
DB 4296 CATCATGTATGCGGCTGTAGATTCGGAAGGCAAGAGCAGACCTGACTACAGAGGCA 4355

QY 1380 TGGAAATTAAGCGGATTAAGCTAGCGGCGCTTCAATGACCACTGGGCGCTTGAAGAGCA 1439
DB 4356 CATCTACATGTATATCTACTCTGTCATCTTCATCATCTTGGGCTCTTCAACCTCAAA 4415

QY 1440 CCTCTTCAAGGCGAGATTAAGCGAGAGCGCA 1471
DB 4416 CCTGTCAATCGGTGTATCATCATGCAACTTCA 4447

RESULT 12
US-09-425-043-2
; Sequence 2, Application US/09425043
; Patent No. 6335172
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.

APPLICANT: HERMAN, RONALD C.
APPLICANT: SANGAMESWARAN, LAKSHMI
TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: JANET PAULINE CLARK
STREET: 3401 HILLYER AVENUE, MS A2-250
CITY: PALO ALTO
STATE: CA
COUNTRY: U.S.A.
ZIP: 94304-1397
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/425,043
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/024,020
FILING DATE: 16-FEB-1998
APPLICATION NUMBER: US 60/039,447
FILING DATE: 26-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: CLARK, JANET P.
REGISTRATION NUMBER: 34,799
REFERENCE/DOCKET NUMBER: R0020B-REG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 852-3097
TELEFAX: (650) 855-5322
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6007 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-425-043-2

Query Match 2.5%; Score 46.4; DB 4; Length 6007;
Best Local Similarity 48.2%; Pred. No. 0.029;
Matches 131; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

DB 1200 GCAGCTGATGGCGAGCAAGGCGTACTGCTGCTGTAACCCGCGGCGAGCGG 1259
4176 GAAGCTCATGAGGAGGCAACAGCAGAGATCCGATGGAGAAGATGTAAGATCAACTTTGA 4235

QY 1260 CTATAGCGAAGACTGCGCGCTCGGCTGGAGAGAGACTGGCTTGAGAGACTTTGAGGA 1319
DB 4236 CATATGTCGAGGAGGAGGACTGCGCTTCTTCAAGTGGCAACCTTCAAAAGCTGGATGGA 4295

QY 1320 CATATGAGACGGCATGAGAGAGTTCTTCAAGCTGGAACCGCAGCGGAGAGGCGCGT 1379
DB 4296 CATCATGTATGCGGCTGTAGATTCGGAAGGCAAGAGCAGACCTGACTACAGAGGCA 4355

QY 1380 TGGAAATTAAGCGGATTAAGCTAGCGGCGCTTCAATGACCACTGGGCGCTTGAAGAGCA 1439
DB 4356 CATCTACATGTATATCTACTCTGTCATCTTCATCATCTTGGGCTCTTCAACCTCAAA 4415

QY 1440 CCTCTTCAAGGCGAGATTAAGCGAGAGCGCA 1471
DB 4416 CCTGTCAATCGGTGTATCATCATGCAACTTCA 4447

RESULT 13
US-09-024-020B-7
; Sequence 7, Application US/09024020B
; Patent No. 6030810
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.

APPLICANT: DIETRICH, PAUL S.
APPLICANT: FISH, LINDA M.
APPLICANT: HERRAN, RONALD C.
APPLICANT: SANGAMESWARAN, LAKSHMI
TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: JANET PAULINE CLARK
STREET: 3401 HILLVIEW AVENUE, MS A2-250
CITY: PALO ALTO
STATE: CA
COUNTRY: U.S.A.
ZIP: 94304-1397
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,020B
FILING DATE: 16-FEB-1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,447
FILING DATE: 26-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: CLARK, JANET P.
REGISTRATION NUMBER: 34,799
REFERENCE/DOCKET NUMBER: R0020B-REG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 852-3097
TELEFAX: (650) 855-5322
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 6556 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-024-020B-7

Query Match 2.5%; Score 46.4; DB 3; Length 6556;
Best Local Similarity 48.2%; Pred. No. 0.03;
Matches 131; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

DB 1200 GCAGCTGATGCGCAGCAAGGCGCTACTGCTGCTGTAACCCGCGCGCAGCGCAGCG 1259
1260 CTATACGGAAGACTTGGCGCTCCGCTCTGAGAGAGACTGGCTTGAGAGACTTTGAGA 1319
DB 4353 CAATGTCGAGGAGGAGGACTGCGCTCTCTTCAAGTGCACCTTCAAGGCTGAGATGGA 4412
QY 1320 CATATGAAAGCGCATGAGAGAGTCTTCAAGCTCAACGCGAGCGCAGAGGAGCGCGT 1379
DB 4413 CATCATGTAATGCGCGCTGATGATCCGAAAGCCAGACGACGACCTGATACGAGGCGAA 4472
QY 1380 TGGAAATPACGGGCGATAGCTAGCGGCGCTCATGACCACTGGCGCTTGACTCAGAGCGA 1439
DB 4473 CATCTCAAGTACATCTACTCTGTCATCTTCATCATCTTGGCTCTTCAACCTCAA 4532
QY 1440 CCTCTCAAGGAGCAAGTAAGCAGAGCGCA 1471
DB 4533 CCGTTCATCGGCTGTCATCATCGCAACTTCA 4564

RESULT 14
US-09-425-043-7
; Sequence 7, Application US/09425043
; Patent No. 6335172
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.

APPLICANT: DIETRICH, PAUL S.
APPLICANT: FISH, LINDA M.
APPLICANT: HERRAN, RONALD C.
APPLICANT: SANGAMESWARAN, LAKSHMI
TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: JANET PAULINE CLARK
STREET: 3401 HILLVIEW AVENUE, MS A2-250
CITY: PALO ALTO
STATE: CA
COUNTRY: U.S.A.
ZIP: 94304-1397
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/425,043
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/024,020
FILING DATE: 16-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: CLARK, JANET P.
REGISTRATION NUMBER: 34,799
REFERENCE/DOCKET NUMBER: R0020B-REG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 852-3097
TELEFAX: (650) 855-5322
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 6556 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-425-043-7

Query Match 2.5%; Score 46.4; DB 4; Length 6556;
Best Local Similarity 48.2%; Pred. No. 0.03;
Matches 131; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

DB 1200 GCAGCTGATGCGCAGCAAGGCGCTACTGCTGCTGTAACCCGCGCGCAGCGCAGCG 1259
1260 CTATACGGAAGACTTGGCGCTCCGCTCTGAGAGAGACTGGCTTGAGAGACTTTGAGA 1319
DB 4353 CAATGTCGAGGAGGAGGACTGCGCTCTCTTCAAGTGCACCTTCAAGGCTGAGATGGA 4412
QY 1320 CATATGAAAGCGCATGAGAGAGTCTTCAAGCTCAACGCGAGCGCAGAGGAGCGCGT 1379
DB 4413 CATCATGTAATGCGCGCTGATGATCCGAAAGCCAGACGACGACCTGATACGAGGCGAA 4472
QY 1380 TGGAAATPACGGGCGATAGCTAGCGGCGCTCATGACCACTGGCGCTTGACTCAGAGCGA 1439
DB 4473 CATCTCAAGTACATCTACTCTGTCATCTTCATCATCTTGGCTCTTCAACCTCAA 4532
QY 1440 CCTCTCAAGGAGCAAGTAAGCAGAGCGCA 1471
DB 4533 CCGTTCATCGGCTGTCATCATCGCAACTTCA 4564

RESULT 15
US-09-024-020B-43
; Sequence 43, Application US/09024020B
; Patent No. 6030810

GENERAL INFORMATION:
APPLICANT: DELGADO, STEPHEN G.
APPLICANT: DIETRICH, PAUL S.
APPLICANT: FISH, LINDA M.
APPLICANT: HERMAN, RONALD C.
APPLICANT: SANGAMESWARAN, LAKSHMI
TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: JANET PAULINE CLARK
STREET: 3401 HILLVIEW AVENUE, MS A2-250
CITY: PALO ALTO
STATE: CA
COUNTRY: U.S.A.
ZIP: 94304-1397
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,020B
FILING DATE: 16-FEB-1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,447
FILING DATE: 26-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: CLARK, JANET P.
REGISTRATION NUMBER: 34,799
REFERENCE/DOCKET NUMBER: R0020B-REG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 852-3097
TELEFAX: (650) 855-5322
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 6586 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-024-020B-43

Query Match 2.5%; Score 46.4; DB 3; Length 6586;
Best Local Similarity 48.2%; Pred. No. 0.03;
Matches 131; Conservative 0; Mismatches 141; Indels 0; Gaps 0;
QY 1200 GCAGCTGATGGCGAGCAAGGCTACTACTGCTGCTCGTAACCCGCGGCGAGCGACGG 1259
DB 4323 GAAGCTCANTGAGGCGCAACGACGAGATCCGATGGAAGATGTCAGATCAACTTTGA 4382
QY 1260 CTATAGCGAAGACTTCGCGCTCCGCTCTGAGAGAGACTGGCTTGAGACTTGAGGA 1319
DB 4383 CAATGTGCGAGGAGGATACCTGCGCTTCTCAAGTGGCAACCTTCAAGGCTGGATGGA 4442
QY 1320 CATATGAAGCGCATCGAGAGATTCTTCAAGCTCGAACCGCAGACGCGGCGGCGT 1379
DB 4443 CATCATGTATGCGGCTGTAGATTCCGGAAGCGCAGAGAGAGAGAGAGAGAGAG 4502
QY 1380 TGGATTAAGCGGCGCATTAAGCGGCGCTTCAAGCTCGAACCTGAGGCGCTGACTCAGAGCGA 1439
DB 4503 CATCTACATGTACATCTACTTGTGATCTTCAATCATCTTGGGCTCTTCTTCAACCTCA 4562
QY 1440 CCTCTCAAGCGAGGAATTAAGCGAGAGCGCA 1471
DB 4563 CCTGTCAATCGGCTGTCATCATCGACAACTTCA 4594

Search completed: July 15, 2003, 15:25:38
Job time : 98 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2003, 10:00:38 ; Search time 433 Seconds

(without alignments)
9720.519 Million cell updates/sec

Title: US-09-966-803-1

Perfect score: 1869

Sequence: 1 ATGACCGCATCGATGAGAA.....TCAGGCGGATGGCACTGA 1869

Scoring table: IDENTITY_NTC

Gapop 10.0 , Gapept 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1869	100.0	1869	AAV12887	Thermococcus amida
2	1078.2	57.7	349980	AAH41226	Pyrococcus abyssi
3	150.8	8.1	265118	AAH41227	Pyrococcus abyssi
4	149.8	8.0	1971	AAK07301	Bacillus subtilis
5	140.8	7.5	1958	AAK72933	Bacillus lichenif
6	120.4	6.4	1896	AAV33582	Pyrococcus horik
7	76.6	4.1	1032	ABO86198	Lactobacillus rha
8	66.2	3.5	477	ABK77575	Bacillus clausii
9	58.4	3.1	1836	AAK60975	P. putida KT2440-a

10	55	2.9	783	23	AA554375	Pseudomonas aerugi
11	51.6	2.8	659	21	AAK13721	Aspergillus oryzae
12	48.2	2.6	1598	14	AAQ40461	E. coli mdh promote
13	47.4	2.5	1062	21	AAI2969	DNA encoding Cellu
14	47	2.5	1785	22	ABA47739	Human breast cell
15	47	2.5	1785	22	ABA56533	Human foetal liver
16	47	2.5	1785	22	ABA32720	Probe #1186 for g
17	47	2.5	1785	22	AAK14036	Human brain expres
18	47	2.5	1785	22	AAK39782	Human bone marrow
19	47	2.5	1785	22	AAI20550	Probe #10523 for g
20	47	2.5	1785	22	AAI5799	Probe #14485 used t
21	47	2.5	1785	22	AAI06289	Human genome-deriv
22	47	2.5	1785	24	AB513877	Human CDNA encodin
23	47	2.5	2428	24	ABK34868	Human wml DNa. Ho
24	47	2.5	3628	21	AAZ88923	Human WFS1 mutant
25	47	2.5	3688	21	AAA08953	Human WFS1 mutant
26	47	2.5	3688	21	AAA08955	Human WFS1 mutant
27	47	2.5	3688	21	AAA08956	Human WFS1 mutant
28	47	2.5	3688	21	AAA08957	Human WFS1 mutant
29	47	2.5	3688	21	AAA08958	Human WFS1 mutant
30	47	2.5	3688	21	AAA09000	Human WFS1 polymor
31	47	2.5	3688	21	AAA09001	Human WFS1 polymor
32	47	2.5	3688	21	AAA09002	Human WFS1 polymor
33	47	2.5	3688	21	AAA09003	Human WFS1 polymor
34	47	2.5	3688	21	AAA09004	Human WFS1 polymor
35	47	2.5	3688	21	AAA09005	Human WFS1 polymor
36	47	2.5	3688	21	AAA09007	Human WFS1 polymor
37	47	2.5	3688	21	AAA09008	Human WFS1 polymor
38	47	2.5	3688	21	AAA09009	Human WFS1 polymor
39	47	2.5	3688	21	AAA09010	Human WFS1 polymor
40	47	2.5	3688	21	AAA09011	Human WFS1 polymor
41	47	2.5	3688	21	AAZ88922	Human WFS1 variant
42	47	2.5	67312	21	AAZ88924	Human DNA sequence
43	47	2.5	123219	23	AAH88703	Human DNA sequence
44	47	2.5	464	20	AAV87011	EST clone AY93. H
45	46.4	2.5	464	20	AAV87011	

ALIGNMENTS

RESULT 1	AAV12887	standard; DNA: 1869 BP.
ID	AAV12887	
XX	AAV12887	
AC	XX	
XX	XX	
DT	22-JUN-1998	(first entry)
XX	XX	
DE	Thermococcus amida	gene.
XX	XX	
KW	Amidase; thermostable enzyme; optically active compound;	
KM	L-amino acid; peptide; peptidomimetic; archaeobacterium; ss.	
OS	Thermococcus sp. strain GU515.	
XX	XX	
FX	Key	Location/Qualifiers
FT	CDS	1..1869
FT		/tag= a
FT		/transl_except= (pos:1228..1230, aa:Val)
FT		/transl_except= (pos:1231..1233, aa:Val)
XX	XX	
PN	MO9748794-A1.	
XX	XX	
PD	24-DEC-1997.	
XX	XX	
PF	17-JUN-1997.	97WO-US09319.
XX	XX	
PR	17-JUN-1996.	96US-0664646.
XX	XX	
PA	(RECO-) RECOMBINANT BIOCATALYSIS INC.	
XX	XX	
PI	Murphy D, Reid JC, Robertson D;	

XX WPI: 1998-063135/06.
DR P-PSDB; AAW41248.
XX Nucleic acid encoding heat-stable amidase from *Thermococcus* GUSL5 -
PT useful for removing N-terminal amino acids from synthetic peptide(s)
PT and peptidomimetic(s)
XX
PS Claim 2; Page 39-41; 56pp; English.
XX This novel isolated nucleic acid encodes a 622 amino acid
CC thermostable amidase (see AAW41248) of the archaeobacterium
CC *Thermococcus* GUSL5. It was isolated from a genomic library of
CC GUSL5 by screening for amidase activity, and subjecting DNA from a
CC positive clone to PCR amplification (see AAW1288-89). The nucleic
CC acid can be used in the production of amidase in host cells, and to
CC identify related sequences encoding similar enzymes. The encoded
CC amidase is used to remove Arg, Phe or Met from the N-terminus of
CC synthetic peptides or peptidomimetics (claimed). Removal of the
CC N-terminal residue can be done even in the presence of a more
CC reactive ester bond (very difficult to achieve non-enzymatically).
CC The amidase is selective for L-amino acids and can therefore be
CC used to produce optically active compounds. The protein tolerates
CC temperatures up to at least 70 degC and high concentrations of
CC organic solvent (e.g. over 40% dimethyl sulphoxide) so can cleave
CC bonds that are normally resistant.
XX
SO Sequence 1869 BP; 476 A; 486 C; 567 G; 340 T; 0 other;
Query Match 100.0%; Score 1869; DB 19; Length 1869;
Best Local Similarity 100.0%; Pred. No.: 0;
Matches 1869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACGGGATCGAATGGAACACGAGACCTTTCTAAGTCCGCTACCGGGCGACCG 60
DB 1 ATGACGGGATCGAATGGAACACGAGACCTTTCTAAGTCCGCTACCGGGCGACCG 60
QY 61 AGGATACGGGGAATTAATCGCGTACACCTCTGACGAAAGGCCAATGAAAGCAAC 120
DB 61 AGGATACGGGGAATTAATCGCGTACACCTCTGACGAAAGGCCAATGAAAGCAAC 120
QY 121 TAGAAGACGCGTGTGTGTGAAGACCTTGAACGGGCTCAAGGGCTTCATCGAAC 180
DB 121 TAGAAGACGCGTGTGTGTGAAGACCTTGAACGGGCTCAAGGGCTTCATCGAAC 180
QY 181 GCCTCAATGCCGAGATTTTGCAGACGGCAGAAAGCTGCTTCAACCTCTTAAACG 240
DB 181 GCCTCAATGCCGAGATTTTGCAGACGGCAGAAAGCTGCTTCAACCTCTTAAACG 240
QY 241 GAGAGAGAGAGACCGAGATATGGTGGCCGATATCCAGACCTGAGCGCCAAAGATC 300
DB 241 GAGAGAGAGAGACCGAGATATGGTGGCCGATATCCAGACCTGAGCGCCAAAGATC 300
QY 301 CTCTCAATCAAAAACGCTCCGCTGATGACATGGAAGACATTCAGAGAACTTAAT 360
DB 301 CTCTCAATCAAAAACGCTCCGCTGATGACATGGAAGACATTCAGAGAACTTAAT 360
QY 361 GTGGGCTTCAAGAGAGAGAGATGAGAGACTTGTCTTTGACAGCAGCTCCGGCTGG 420
DB 361 GTGGGCTTCAAGAGAGAGAGATGAGAGACTTGTCTTTGACAGCAGCTCCGGCTGG 420
QY 421 TTCGACATATATGAGATTTTGAATGAGAGAGAGACGATTCGGGTTCTTGAACATG 480
DB 421 TTCGACATATATGAGATTTTGAATGAGAGAGAGACGATTCGGGTTCTTGAACATG 480
QY 481 GCCGAGAGATATATCGAGAGCTTCGAGAAACCGAGTTTTCGAGTGGCTTGGCAGGC 540
DB 481 GCCGAGAGATATATCGAGAGCTTCGAGAAACCGAGTTTTCGAGTGGCTTGGCAGGC 540
QY 541 GATGCAATAGTTGTGAACGTCGCCGACCGGAGGAGAGCAAGCTGCTTCAAGATTC 600
DB 541 GATGCAATAGTTGTGAACGTCGCCGACCGGAGGAGAGCAAGCTGCTTCAAGATTC 600

QY 601 TAGACATAGTCTCTATGGAAGAGCGGGAGAGAAAGCAAGCTCTTGCAGAGGCTTC 660
DB 601 TAGACATAGTCTCTATGGAAGAGCGGGAGAGAAAGCAAGCTCTTGCAGAGGCTTC 660
QY 661 GAGCGGTGACTCCGACGGAAGAGAAATCTCTGAGGGGCAAGAAAAAGCGGTTTC 720
DB 661 GAGCGGTGACTCCGACGGAAGAGAAATCTCTGAGGGGCAAGAAAAAGCGGTTTC 720
QY 721 ATCAGCGAGCAGACATGCTGTACCTCTGTGGAGACGGGACCTTAACCCATTCAGAGGC 780
DB 721 ATCAGCGAGCAGACATGCTGTACCTCTGTGGAGACGGGACCTTAACCCATTCAGAGGC 780
QY 781 CCCTGACGCTCTGGGAGAGCAAGCTCAGGAGGAAAGGTACTTCTCTACTCCAGAT 840
DB 781 CCCTGACGCTCTGGGAGAGCAAGCTCAGGAGGAAAGGTACTTCTCTACTCCAGAT 840
QY 841 GCGGGAGGGTAAACCTCTGCTGTGGAGCGGAGCGGAGCGGTTGTACCGGCGAC 900
DB 841 GCGGGAGGGTAAACCTCTGCTGTGGAGCGGAGCGGAGCGGTTGTACCGGCGAC 900
QY 901 CACTGATTTACGGGGCTGTGACGTCAGCGATGGCAAGCATGTCTCTCATGACCGCC 960
DB 901 CACTGATTTACGGGGCTGTGACGTCAGCGATGGCAAGCATGTCTCTCATGACCGCC 960
QY 961 ACAGGATAGCGAGCTCTACCTCTACGACGCGGACGCTGAAACAGGTCACCGAATCAAC 1020
DB 961 ACAGGATAGCGAGCTCTACCTCTACGACGCGGACGCTGAAACAGGTCACCGAATCAAC 1020
QY 1021 GGGCCGATATTCAGGAAGCTCAAGACCTTGCAGCGGACGCTCCGCTTCAAGACAA 1080
DB 1021 GGGCCGATATTCAGGAAGCTCAAGACCTTGCAGCGGACGCTCCGCTTCAAGACAA 1080
QY 1081 GACCTGAGATTAAGCGGCTGTGTACCTGAGCGGAGGTTAAAGAGAGAAAGCGCGG 1140
DB 1081 GACCTGAGATTAAGCGGCTGTGTACCTGAGCGGAGGTTAAAGAGAGAAAGCGCGG 1140
QY 1141 ATAGTCTCTCCACGCGGCGGCGAAGGGCATGTACGACACCGCTTCTTACGAGAT 1200
DB 1141 ATAGTCTCTCCACGCGGCGGCGAAGGGCATGTACGACACCGCTTCTTACGAGAT 1200
QY 1201 CAGCTGATGGCGAGCAAGGGCTACTACTGTCTGTGTGTAACCGGCGGACGCGG 1260
DB 1201 CAGCTGATGGCGAGCAAGGGCTACTACTGTCTGTGTGTAACCGGCGGACGCGG 1260
QY 1261 TATAGCAAGACTTCCGCGCTCCGCTCTGAGAGAGACTGCTTGAAGACTTTGAGAC 1320
DB 1261 TATAGCAAGACTTCCGCGCTCCGCTCTGAGAGAGACTGCTTGAAGACTTTGAGAC 1320
QY 1321 ATATGAACGGCATCGAGAGTCTTCAAGCTCGAACCAGGCGCAGAGAGCGCGTT 1380
DB 1321 ATATGAACGGCATCGAGAGTCTTCAAGCTCGAACCAGGCGCAGAGAGCGCGTT 1380
QY 1381 GGAATTAAGGGGCAATGAGCTACGGGCTTATGACCAACCTGGGCTTGTACTGAGCGAC 1440
DB 1381 GGAATTAAGGGGCAATGAGCTACGGGCTTATGACCAACCTGGGCTTGTACTGAGCGAC 1440
QY 1441 CTCTTCAAGCAGGAATTAAGCAGAAAGGCAATGAGTACTGTGCTACACAGCTTAC 1500
DB 1441 CTCTTCAAGCAGGAATTAAGCAGAAAGGCAATGAGTACTGTGCTACACAGCTTAC 1500
QY 1501 TCGGACATATAGGCTCTGTGACAGCTCGAGCTATCGGGCAATCCGTTAGAGAGAG 1560
DB 1501 TCGGACATATAGGCTCTGTGACAGCTCGAGCTATCGGGCAATCCGTTAGAGAGAG 1560
QY 1561 AACCTCAGGAAGCTCAGCCCGCTGTCTTACGCTCAGAGCTGAAAGGCGCGATCTCT 1620
DB 1561 AACCTCAGGAAGCTCAGCCCGCTGTCTTACGCTCAGAGCTGAAAGGCGCGATCTCT 1620
QY 1621 ATCCACTCGCTTGAAGACTACCGCTGTCCGCTGACAGAGCCTTATGTTTCAACAGTG 1680
DB 1621 ATCCACTCGCTTGAAGACTACCGCTGTCCGCTGACAGAGCCTTATGTTTCAACAGTG 1680
QY 1681 CTCAGAGCATATGGCAAGAGACCTTACATATGCAATATTCAGCGCGGCGCCACGCGCAC 1740

Db	1681	CTCAAGGACATGGGCAAGAGCCTACATAGCATATCAAGCGGGGGCCACAGGCCAC	1740
Qy	1741	AGCGTCGGCGGAACCCGAGGCACAGGCCAAGGCGCTACAGGCTCTTCATAGAGTTCTTC	1800
Db	1741	AGCGTCCGCGAAGCCCGAGGCACAGGCCGAAAGCGCTACAGGCTCTTCATAGAGTTCTTC	1800
Qy	1801	GAGCGCAAGCTCAAGAAGTACGAGAGGCGCTTTAGGTAGAGATGATCTCAAGGGGAAT	1860
Db	1801	GAGCGCAAGCTCAAGAAGTACGAGAGGCGCTTTAGGTAGAGATGATCTCAAGGGGAAT	1860
Qy	1861	GGGAACCTGA	1869
Db	1861	GGGAACCTGA	1869

Seq	Sequence	349980 BP; 98084 A; 80447 C; 77665 G; 93784 T; 0 other;
XX	Query Match	57.7%; Score 1078.2; DB 22; Length 349980;
	Best Local Similarity	74.9%; Pred. No. 5,9e-260;
	Matches 1405; Conservative	0; Mismatches 458; Indels 12; Gaps 4;
QY	1 ATGACCGGCGATCGAATGTAAGCAACCGAGACCTTTTCAATGTTGCCCTTACCTCGGGCGACCCG 60	
DB	206185 ATGACCTCTATCGAATGTAATGTAAGCAAGCTTTTACAAAGTTGCCCTTACCTTACCGACCC 206128	
QY	61 AGGATACGGGGAAACCTTAATCGCGTACACCTGACGAGAGCCCAATGAAAGCAACAAAG 120	
DB	206125 AGGATAAAGGGGAAGCAACAAATAGGTGTCAGTTCTTACCAAGGCGAACCTTGGACAAACAAAG 206066	
QY	121 TACGAGAGCAGCGTTTGTGTTGAACCTTGAACGGGCTCAAGCGCTTCATCGAAGAC 180	
DB	206065 TACGAGAGCAGCATATCTTCTCGAAGACCTCGAAGATGGGAGCGCAAGGGTTCATATGAAGAC 206008	
QY	181 GCCCTAATGCCGAGATTTTCGCCAGACGGCAGAAAGCTGGCCCTTCAACCTGCTTAAACGAG 240	
DB	206005 GCTTCAATGCCCTTATAGATTTCTCCAGATGGAAGAAAGATAGCTTCTCATGAGTTCAACGAG 205948	
QY	241 GAGAAGAAAGAGACCGAGATATGGGTGGCCGATATTCAGAACCTTGAGCCGCAAGAAAGTC 300	
DB	205945 GAGAAGAAAGATGCTCTCAGATATGGGTGGCGGACATGAAACCTTAAAGCCGAAAGAGTT 205886	
QY	301 CTCTCAACTAAAAAGTCGCGCTCGATATGCGAGTGGAAAGCATGATTCAGAGACACTCTTAAGTT 360	
DB	205885 CTGAGGCGCCAAAGAAATTAAGTCCCTGCAATGGAATGACAGACTCAAGAGAGTTGTATGT 205826	
QY	361 GTGCGCTTCAAGAGAGGAGGAGCATGAGACTTCGCTTTTACAGACAGACGCTCCCGATCTGG 420	
DB	205825 ATAGGCTTCAAGAGAGGAGGAGCATGAAGATTCATATGAGAGATGATCCAGCTTGG 205766	
QY	421 TTGCGCAATATGGGATATCTTTGATGAGAGAAAGACAGACTCTGSGTCTTGACACTGAG 480	
DB	205765 TTGCGCAATATGGGATATCTTTGACAGGCGCGAAAGACACTTGTGATTAATATATCTGAG 205706	
QY	481 GCCGAGAGATTAATGAGCAGATTGAGAAAGCCGAGGTTTTCAGTGGACCTCGGCGACGAG 540	
DB	205705 GCCGAGAGAGTTATATGAGCAGTTCGAGAAAGCCAAAGTTTTCGACGGAATTTGGCAGCGG 205646	
QY	541 GATGCGATATGTTGTGAACGTCCCGACCGCGAGGGAGCAGACTGCGCTGTTCAGATTG 600	
DB	205645 GATTCGATATGTTGTGAACGTCCCGCTACAG--GGAACCAATACCAAGTACTTCAAGTAC 205589	
QY	601 TACGACATATGCTCTATGGAAGAGCGGGGAGAAAGAAAGCTTTGAGAGGGTCTTCCTTC 660	
DB	205588 TGGAAACATCTACCTCTGGAAGAGATGAGAGGAAAGAAAGCTATTTGAGAGAGGTTCTTTC 205529	
QY	661 GAGCGGTTGACTCCGACGGAAGAAAGAAATCTCTGAGGGGCGMAAAAAAAGGCGGTTG 720	
DB	205528 CACCTTATAGACTCTGACGGAAGAAATTTCTACTCTACGAAAAAGCCGAGAAAGAGTAC 205469	
QY	721 ATCAGCGAGCAGACTGGCTGATCTCGTGGAGAGCGCGAGCTTAA---ACGATCTACAG 777	
DB	205468 ATGAGCAGCAGATTAAGCTCTTACACTTACGACGCGGAGGTTAAGGAATATGGAATTCG 205409	
QY	778 GGGCCGCTCGACGTCTGGGAAGCCAAAGCTCAAGAAAGAAAGTCTACTTCTCTCACTCA 837	
DB	205408 ATAGACAGAGAAAGCTGCTCAAGCTATGATCAAGGATGGGAAGGTGTACTTCACTATAC 205349	
QY	838 GATCGGGGAGGAGGTAACCTCTGCTGTGGGACGGGAAGGCCGAGCGTGTGTTACCGGC 897	
DB	205348 GAGGAGGGAAGCCGTTAACTATATCTCTGGGATGGGATATCAAGAGATAGCTTAAGGGC 205289	
QY	898 GACCACTGGAATTTACGGGCTTGAGCTCAGCATGGCAAGATGCTCTCATCATGACC 957	
DB	205288 AAGCACTGGAATATAGGGGTTGCAAGT---TACAGAAAGGTGTGCTACTCTACAGAAACT 205232	
QY	958 GCCACGAGGATAGCGGACTCTACTCTTACGAGGGCGAGCTGAAAACAGTCAACCGAATAC 1017	

Db 1771 CTTTGATACGCGATGGCGAGCGGGATGACCGGATGCCGATCGAGCGCGGAGCAGCTG 1830

QY 1872 TACACAGCTGCTCAAGGACATGGGCGAAGGAAGCCTACATAGCGATATTCAAGCGCGGCCCC 1731

Db 1831 TTTATCCGCTCTGAAAAAATGGGCGAAGGAACCAAGCTTCTCCGTTTCCGATATGCATCG 1890

QY 1732 CAGCGCCACAGACGTCGCGGGAGGCCGAGGCGACAGCGCGGAAGGCGCTACAGGCTCTGTATA 1791

Db 1891 CACAAATTTATCAGCGACCGGACACCCCAAGACAGCGGATCAAGCGCCTGAAATTTATATCAGC 1950

QY 1792 GAGTCTTCGAGCGCAAGCTC 1812

Db 1951 TCATGGTTGATCAACATCTC 1971

RESULT 5
 ABR72933
 ID ABR72933 standard; DNA; 1958 BP.
 AC ABR72933;
 DT 13-AUG-2002 (first entry)
 XX
 DE Bacillus licheniformis genomic sequence tag (GST) #224.
 XX
 KW Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX
 OS Bacillus licheniformis.
 PN WO200229113-A2.
 PD 11-APR-2002.
 PF 05-OCT-2001; 2001WO-US31437.
 PR 06-OCT-2000; 2000US-0680598.
 PR 27-MAR-2001; 2001US-279526P.
 PA (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 PI Berka R, Clausen IG;
 PI WPI; 2002-416684/44.
 DR
 XX
 PT Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second
 PT Bacillus cells, by using substrate containing Bacillus genomic
 PT sequenced tag array -
 PS
 PS Claim 4; SEQ ID NO 224; 200pp; English.
 XX
 CC The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at
XX ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 1958 BP; 557 A; 462 C; 518 G; 421 T; 0 other;

Query Match	7.5%;	Score 140.8;	DB 24;	Length 1958
Best Local Similarity	51.8%;	Pred. No. 2.6e-25;		
Matches 344;	Conservative	0;	Mismatches 317;	Indels 3;

QY	1115	AGGTTAAAGAGAAAGGCCCGGTGATAGCTTCGTCACAGGGGGGCCCAAGGGCATGT	1174
Db	1259	ATGAAAAGGAAACAGCTATCCCGTCACTTCGTGAAGTGCACGGGGACCTCACCCCATGT	1318
QY	1175	ACGACACACCCGCTTCGTACAGATGCAGCTGATGGCGACCAAGGCGCTACTACTGCTGCT	1234
Db	1319	ATCGGAATGCGTTATTTTCATGAAATTTCAATGATGCTGCGCCGCAAGAGGACCGCTCGCTAT	1378
QY	1235	TCGGAAACCCGGCGGAGAGGAGCGGCTATAGCGAAGACCTTCGGCGCTCCGGGTCTGGAGA	1294
Db	1379	ATGTTAAACCCGGGGGAAAGCCAGCGCTACGGACAGGATTTTGTCAACCGGGGTGAGGGCG	1438
QY	1295	GGACTGGCTTGGAGAGCTTTGAGGACATATGAAACGGCATCGAAGAGTTCTTCAAGCTCG	1354
Db	1439	ACTACGGGGGTGGAGATTTTAAAGATGATTAATGGCGGCTGTGTGACCACTGCTTGAACATT	1498
QY	1355	AACCGCAGGCCGACAGGAGGAGCGGCTTGGAAATACGGGCATTAAGTACGGGGGCTTTCATGA	1414
Db	1499	ATGATTTTCGTGATCAAGAAAGGCTCGGACATACCGGAGGAAAGCTACGGGGGCTTTATGA	1558
QY	1415	CCAACCTGGGCGCTTGACTCAGAGCAGCACTCTTCAAGGACAGAAATPAGGACAGAACGGCATAA	1474
Db	1559	CAAACTGGGCGTGTGGGCAACAGAAAGGCGTTCAAGGGCGGAGTGCACACAGAGGTCCATTT	1618
QY	1475	GCTACTGGCTCACAGCTAGCGCTTCTTCGGACATAGGCGCTGTGTTACAGAGTGCAGGTCA	1534
Db	1619	CCAACCTGGATCAAGTTTTCAGGGGTAAAGCGCAATCGGCTATTTTTCACACAGACGGGAGC	1678
QY	1535	TCGGGGCCAATCCGCTT---AGAGAAAGAGAACTCAGAAAGCTCAGCGCGGCTGTTCACG	1591
Db	1679	TCGGTGCAGACCTTTTGGAAAGATTCACAGCAAACTGTGGAGCATTTCCCGCTGCATAAATAGC	1738
QY	1592	CTCAGAAAGTGAAAGCGCCGATACTCTTAATCCACTCGCTTGAGAGACTACCGCTGTCCGC	1651
Db	1739	CGGACAAAGGTGGAGAACCCCTCTCTTCTCAATTCTGTGCTGGAGAACGGGAGACAGGCGCCGA	1798
QY	1652	TCGACCAAGAGCGCTTATGTTCTACACAGTGGCTCAAGAGCATGGGCAAGAAAGCCATCAATAG	1711
Db	1799	TTGTAACAGCGGAAACAGCTGTTCCACGGCGCTGAAAATAATGGGAAAGAAATATAACTCG	1858
QY	1712	CGATATTCAAGCGCGGCGCCACAGGCGACAAAGCGTCCGCGGAAGCCGAGACACAGGCCGA	1771
Db	1859	TCAGATTTCCGAAGCGCGTCAACAGATTTATCAAGAAACGGGACATCCGAAGACGGGATCA	1918
QY	1772	AGCG 1775	
Db	1919	GGCG 1922	

	RESULT 6
AAV33582	
ID	AAV33582 standard; DNM; 1896 BP.
XX	
AC	AAV33582;
XX	
DT	29-DEC-1998 (first entry)
XX	
DE	pyrococcus horikoshi acylpeptide hydrolase encoding DNA sequence
XX	
KW	Pyrococcus horikoshi acylpeptide hydrolase; ss.
XX	
DS	Pyrococcus horikoshi.

XX	Key	Location/Qualifiers
FH		1.1896
FT	CDS	/tag-a
FT		/product= "Pyrococcus horikoshi acylpeptide hydrolase"
FT		/note= "CDS does not contain a stop codon"
XX		JPI0210977-A.
PD		11-AUG-1998.
XX		31-JAN-1997; 97JP-0018381.
XX		31-JAN-1997; 97JP-0018381.
XX		(AGEN) AGENCY OF IND SCI & TECHNOLOGY.
PA		WP1; 1998-488369/42.
DR		P-PSDB; AAM70508.
XX		A heat-resistant acylpeptide hydrolase and a gene coding it - useful
PT		for hydrolysing the C terminl of proteins at high temperature
XX		Example 5; Pages 6-7; 8pp; Japanese.
XX		The present sequence represents a Pyrococcus horikoshi acylpeptide
CC		hydrolase encoding DNA sequence. The enzyme has the following
CC		properties: (a) It hydrolyses an acylpeptide; (b) an optimum temperature
CC		of 90-95 degrees Centigrade; (c) an optimum pH of 5.0-6.0; (d) no loss
CC		of activity when heated to 95 degrees Centigrade for 3 hours at pH 7.5;
CC		and (e) a molecular weight of 60 kDa. The enzyme is claimed useful
CC		for hydrolysing the amino end of an acylated protein and a peptide at
SQ		high temperatures.
XX		Sequence 1896 BP; 614 A; 358 C; 491 G; 433 T; 0 other;
Query Match	6.4%; Score 120.4; DB 19; Length 1896;	
Best Local Similarity	50.3%; Pred. No.3.3e-20;	
Matches 351; Conservative	0; Mismatches 341; Indels 6; Gaps 2;	
DY	1120 AAAGAGAAGAAAGGCCCGGTGATGCTTCGCCACGGCGGGCCGAAGGCATGACGA	1179
DY		
DY	1198 AAAAGAAAAGATGATCCAGCATATTCTAGAGATCCACGGTGTCTTAAACCCTTAGCGGT	1257
DY	1180 CAACGCTTCGTTACGAGATGACAGCTGATGGCGAAGAGGCTACTACTGCTGCTGTG	1239
DY		
DY	1258 TACGCTTTATGACACGAGTTCCACGTTTTTAACCTTAAGAGGCTTCGTGATATTCTCA	1317
DY	1240 AACCGCGCGGCACGACGGCTATAGGAAGAACTTCGCGCTCCGCGTCTGGAAGAGACT	1299
DY		
DY	1318 AATCTTGAGAGGACGATGGCTTACGAGAGAGATGTCGCGATA--TAAAGGACACTAT	1374
DY	1300 GGCTTGGAGACTTTGAGACATPATAAACGGCATCGAGAGTTCTTCAACCTGGAACCG	1359
DY		
DY	1375 GGGGAGAGGGATTACAGAGATTATATGAGAGTATCCATGAAGCATTAAGAGATTGAC	1434
DY	1360 CAGGCCACAGAGGAGCGCGTTGATTAACGGGCTAAAGCTACGGCGGCTTCATGACCAA	1419
DY		
DY	1435 TTCAATAGATGGGAAAGGCTAGAGATTACCGGGGGGTTCCTATGGTGCTTCATGACGAA	1494
DY	1420 TGGGCGCTTGAATCGAGACGACCTCTTCAAGCGAAGATTAAGCGAAGACGATTAAGTAC	1479
DY		
DY	1495 TGGAATGTCGACATCCAAACAGGTTCAAAGCCCGCTGTACAACCAAGATTTCAAT	1554
DY	1480 TGGCTCACAGACTACGCTTCTCGAGATATAGGGGCTGTGATACGACGTGAGGTATCGGG	1539
DY		
DY	1555 TGGAATGACTTCTTCGSGGACACAGGATATAGGTTATTACTTTCCTCAGATCAAAATAGA	1614
DY	1540 CCAATACCGTTAGAGAAC--GAGAACTTCAGAGAGCTCAGCGCGCTTCTCAGCTCAG	1596
DY		
DY	1615 AAAGATCCCTGAGACAACTTGGAAGGTATTGGGAAAAAGGCCCATTTAAGATACGCTCC	1674
DY	1597 AAGCTGAAGCGCGGATACCTCTAATCACTCGCTTAGAGGACTACCGCTGCTGCTGAC	1656

Db	1675	AAAGTGAAGTAACTCCCTCCCTGCTTAATATCAACCTACCGAAGACTACAGGTGTGGCTTCCC	17344
Oy	1657	CAGAGCCCTTATGTTCTACACAGTGTCTCAAGGACATGGGCAAGGAAGCCTACATFAGCGATA	1716
Db	1735	GAGCGATTGCACACCTTCTATCATATCCCTAAATAAATCTGGGGAGAGAGTGGAAFTTGGCAATA	17944
Oy	1717	TTCAAGCGCGGCGCCACAGCGCCACAGCGCTCCGCGAAGCCCGAGGACAGCGCCAGAGCGC	1776
Db	1795	TTCCGAGGAGAAAATCATGACCTAGTATGATCTGGGAAGCCAAAGACAGGCTTAAAGA	1854
Oy	1777	TACAGGCTCTTCATAGAGTTCTTTCGAGCGCAAGCTCAA	1814
Db	1855	CTTGAACTAAATAGCAGATGATGATGAGAAATAGCTTAA	1892
RESULT 7			
ABQ86198	ID	ABQ86198 standard; DNA; 1032 BP.	
XX	AC	ABQ86198;	
XX	DT	09-SEP-2002 (first entry)	
DE	DE	Lactobacillus rhamnosus HN001 polynucleotide SEQ ID NO 14.	
XX	XX	Lactobacillus rhamnosus; strain HN001; 'vulnerary; antilipaeemic; milk;	
KW	KW	immunostimulant; anti-infection; lactose digestion; immune system;	
KW	KW	metabolic activity; nutrition; health; transgenic; lactic acid bacteria;	
KW	KW	genome mapping; gastrointestinal disorder; dairy processing; vaccine;	
KW	KW	fermentation; probiotic; cholesterol; wound healing; gene; ds.	
XX	XX	Lactobacillus rhamnosus.	
OS	OS	WO20024383-A1.	
PN	PN	06-JUN-2002.	
XX	XX	28-NOV-2001; 2001WO-NZ00286.	
XX	XX	28-NOV-2000; 2000US-0724623.	
PR	PR	(GENE-) GENESIS RES & DEV CORP LTD.	
XX	PA	(VIAL-) VIALACTIA BIOSCIENCE NZ LTD.	
XX	PA	Glenn M, Havukkala JU, Lubbers MW, Dekker J;	
PI	PI	WPI; 2002-519588/55.	
XX	DR	P-PSDB; ABP61033.	
PT	PT	Novel isolated Lactobacillus rhamnosus polynucleotides encoding	
PT	PT	polypeptide with anti-infection/lactose digestion modulating activity,	
PT	PT	useful to improve properties of microbes used in milk-derived products	
PT	PT	manufacture	
XX	XX	Claim 2; Page 63; 128pp; English.	
XX	XX	The invention relates to an isolated polynucleotide (I) comprising a	
XX	XX	nucleotide sequence (ABQ86185-ABQ86243) present in Lactobacillus	
CC	CC	rhamnosus strain HN001 that encodes a polypeptide (ABP61020-ABP61060)	
CC	CC	with activity such as enzyme activity; anti-infection activity; lactose	
CC	CC	digestion modulating activity; immune system modulating activity; amino	
CC	CC	acid, lipid, vitamin or carbohydrate metabolic activity; flavour, texture	
CC	CC	or aroma modulating activity; (I) is useful for improving the properties	
CC	CC	of microbes used in the manufacture of milk-derived products and	
CC	CC	probiotic supplements, which involves modulating the polynucleotide	
CC	CC	content or composition of the microbes by transforming the microbes with	
CC	CC	(I). (I) is also useful for identifying an organism (preferably a	
CC	CC	bacterial or yeast cell) or reproductive material or an extract from the	
CC	CC	organism, as having a specific origin. Proteins encoded by (I) are useful	
CC	CC	for modifying the flavour, aroma, texture and/or nutritional and health	
CC	CC	benefits of milk-derived products, which involves adding one or more	
CC	CC	polypeptides to the milk being processed. Genetic constructs comprising	

OS Pseudomonas aeruginosa.
 XX
 PN W0200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen K, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR P-PDB; AA036516.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Claim 27; Seq ID No 8012; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 CC Sequence 783 BP; 183 A; 268 C; 224 G; 108 T; 0 other;
 XX
 SQ
 Query Match 2.9%; Score 55; DB 23; Length 783;
 Best Local Similarity 45.1%; Pred. No. 0.00066;
 Matches 205; Conservative 0; Mismatches 250; Indels 0; Gaps 0;
 QY 882 GCGTGTGTTACCGGCGACATGATTTACGGGCTTGACGATGCAAGCATT 941
 DB 12 GCTTGGCTTTCGCGCTGCGGGGCTCGGCGCTACAGCGGCCGCCGCCGCGT 71
 QY 942 GCTTCATCATATACCGCCGACGAGATAGCGGCTTACCTTACGACGGCGATGAA 1001
 DB 72 CTTACCGCTGCGCGCACCCGGGTGCGCGACGGGAGANTCCTACGCTGTCAGACCGCT 131
 QY 1002 ACAGTTCACCGATTAACAGGGCGGATATTCAGAGAGCTCAAGCTTCGAGCGGAGCA 1061
 DB 132 GCTGGCCAGGAAGGCGTGACCTGAAGATCAAGAGATTCACGATACCTGACGCCGA 191
 QY 1062 CTTCCGCTTACAGCAAGACCTCGAGATAGACGGCTGTACCTACAGCGCGAGGTTAA 1121
 DB 192 CGTGACAGGCTCGGAAAGGCGCTGAGACGCACTTCTTCACACACAGCCGTAACCTGCA 251
 QY 1122 AAGGAGAAAGGCGCCGCTGATCTTTCGTCACGGCGGCGCAAGGCGCATGTACGACA 1181

DB 252 TGAGTTCAACAAGGCCAAGGGCACCGACCTGTGCGCGTACCGGCGTACACATCGAGCC 311
 QY 1182 CCGCTTCGCTTACAGATGACGTATGCGGACCAAGGCGTACTACTCTTCTGTAA 1241
 DB 312 GCTGGGGCGCTTACTGAGCAAGTACAAAGAGCTGACGACTCTCCGGCGCTACCGCT 371
 QY 1242 CCGCGCGGAGGAGGCGCTATAGCGAAGCTTCGCGCTCCGCTCTGAGAGGAGCTGG 1301
 DB 372 GGTGATTTCCCAAGGACGCCACCAAGCGGCGCGCGCTCTCTCTGACAAAGCGCG 431
 QY 1302 CTTGAGGACTTGTAGGACATTAATGAACGGCATCG 1336
 DB 432 GGTGATCAAGCTCAAGGACCAACAAGGACATCACCG 466
 RESULT 11
 AAF13721
 ID AAF13721 standard; cDNA; 659 BP.
 XX
 AC AAF13721;
 XX
 DE 13-MAR-2001 (first entry)
 XX
 DE Aspergillus oryzae EST SEQ ID NO:6244.
 XX
 KW Multiple gene expression: filamentous fungal cell; EST;
 KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
 KW Aspergillus oryzae; Trichoderma reesei; Identification; recombination;
 KW culture condition; environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.
 OS
 XX Aspergillus oryzae.
 XX
 PN W0200056762-A2.
 XX
 PD 28-SEP-2000.
 XX
 PF 22-MAR-2000; 2000WO-US07781.
 XX
 PR 22-MAR-1999; 99US-0273623.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Berka RM, Rey MW, Snustner JR, Kauppinen S, Clausen IG, Olsen PB;
 XX
 DR WPI; 2000-594572/56.
 XX
 XX
 PT Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags -
 XX
 PS Claim 88; Page 2569; 3161pp; English.
 XX
 CC The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring
 CC the global expression of genes from FF cells allows the production
 CC potential of the microorganisms to be improved. New genes may be
 CC discovered, possible functions of unknown open reading frames can be
 CC identified and gene copy number variation and stability can be
 CC monitored. The expression of genes can be used to study how FF cells
 CC adapt to changes in culture conditions, environmental stresses, spore
 CC morphogenesis, recombination, metabolic or catabolic pathway
 CC engineering. Using ESTs provides several advantages over genomic or
 CC random cDNA clones including elimination of redundancy as one spot on an
 CC array equals one gene or open reading frame, and organisation of the
 CC microarrays based on function of the gene products to facilitate

XX	ID	AAAL2989	standard; DNA; 1062 BP.
AC	XX	AAAL2989;	
XX	DT	18-JUL-2000	(first entry)
XX	DE	DNA encoding Cellulomonas fimi xylanase.	
XX	KM	Xylanase; endo-1,4-beta-xylanase; xylan hydrolysis;	
XX	KW	xyloligosaccharide production; ds.	
OS	XX	Cellulomonas fimi.	
XX	FN		
XX	XX		
XX	PD	10-APR-2000.	
XX	XX		
XX	PF	17-MAR-1999; 99JP-0071715.	
XX	PR	17-MAR-1999; 99JP-0071715.	
XX	PA	(NORU) NORIN-SUISANSO. SHOKUHI SOTO KENKYUSHOCH.	
XX	PA	(SEIB-) SEIBUTSUKETI TOKUTEI SANGYO GIJUTSU KENKYU SUTSHIN KIKO.	
XX	DR	WPI: 2000-306509/27.	
XX	DR	P-PSDB: AAY81496.	
XX	XX		
XX	PT	Modified xylanase gene for use in foodstuff industry, has base sequence	
XX	PS	of 1065 nucleotides -	
XX	XX		
XX	XX	Example 1; Page 9-10; 11pp: Japanese.	
CC	CC	The invention relates a novel chimeric xylanase (AAY81494) and to DNA	
CC	CC	encoding it (AAAL2985). The chimeric xylanase comprises residues 1-244	
CC	CC	of Streptomyces olivaceoviridis xylanase and residues 244-354 of	
CC	CC	Cellulomonas fimi xylanase. Xylanase (also known as endo-1,4-beta-	
CC	CC	xylanase) hydrolyses beta-1,4-D-xylan, a component of the hemicellulose	
CC	CC	in plant cell walls, into xyloligosaccharides and xylose. Xylanase is	
CC	CC	used in a range of industrial processes. It is used to produce xylol-	
CC	CC	oligosaccharides from xylan from broad-leaved trees, and is used in	
CC	CC	wood-pulp bleaching to reduce the amount of chlorine required for this	
CC	CC	process. Xyloligosaccharides can be used as ingredients in foodstuffs	
CC	CC	and as water-retaining material in cosmetics. The chimeric xylanase of	
CC	CC	the invention does not generate xylose monomers during the hydrolysis of	
CC	CC	xylan. It is therefore useful for efficient and reliable xylol-	
CC	CC	oligosaccharide production. The present sequence represents DNA	
CC	CC	encoding xylanase from Cellulomonas fimi.	
XX	XX		
XX	XX	Sequence 1062 BP; 181 A; 388 C; 358 G; 135 T; 0 other;	
XX	XX		
XX	XX	Query Match 2.5%; Score 47.4; DB 21; Length 1062;	
XX	XX	Best Local Similarity 43.3%; Pred. No. 0.059;	
XX	XX	Matches 222; Conservative 0; Mismatches 291; Indels 0; Gaps 0;	
QY	829	CTCACTCCAGATGCGGGCGAGGATAACCTCTGGCTTGGAGCGAGGAAGCCGAGCGTGT 888	
DB	67	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT	
QY	889	GTTACCGCGGACCACTGGATTATTACGGGCTTACGTCGACGCGATGGCAAAAGCACTGCTCTC 948	
DB	127	GCGACCAACGCTCAAGAGAGCCGCCGACGCGGCCGCCGCGGACTTGCTTGCGCTGCAC 186	
QY	949	ATCATGACCGCGACGAGATGAGGAGAGCTTACCTCTACGACGCGGAGCACTGAACAGATC 1008	
DB	187	CCCAACCGGCTCTGGGAGGCGCAGTACAAAGGCGATCGCGACAGGAGATTCAACTGCTC 246	
QY	1009	ACCGAATACAAAGGGCGGATTTTACAGGAGCTCAAGACTTCAGACCGAGGCACTTCCGC 1068	

Db	247	GTGGCGGAAGACGGATGAATGGGAGCGCACCGAGGCCCTGGCGAAGAACACTTCTCTTC	306
QY	1069	TTCCAGAGCAAAAGACCTGCAGATAGACGGCTGTACTCAAGCCCGAGGTTAAAGAGAG	1122
Db	307	GGGGCGGGGACACGGGTGCGGAGTACCCGCCACACACGGCAAGGAGGCTGTACGGCCAC	366
QY	1129	AAGGCCCGGCGTATGATCTTGGTTCGACGGGGGGCGAAGGCCATGTATGACACCGCTTC	1188
Db	367	ACGCTGTATWGCACTGCGACGTGCCGACGTGGGGAGAACTCAACGGCTCCGGCTTC	426
QY	1189	GTCCTACGATGCACGCTGATGGCGAGCAAGGGCTACTACTGCTTCGTGAACCCGCGC	1241
Db	427	GAGAGCGCATGGTCAACACCTGACAGAAAGTCCGCCACACTTCGAGGGCAAGATCGGG	486
QY	1249	GGCAGCGACGGCTTATAGCGAAGACTTCGCGCTCCGCGTCTGGAGAGGACTGGCTTGGAG	1301
Db	487	TCTGGGAGGTGCTCAACAGGCGCTTCCGCCACGGCGGGCGGCCGCCGCGAGGACTCGGG	546
QY	1309	GACTTGAGACATAATGAAAGGGCATGAGAGG	1341
Db	547	TTCCAGCAAGAGCTCGGCACACGGCTACATCGAG	579

RESULT 14

ID ABA47739 standard; DNA; 1785 BP.

AC ABA47739;

DT 01-FEB-2002 (first entry)

DE Human breast cell single exon nucleic acid probe #6434.

KW Human; microarray; single exon probe; gene expression; breast;

KW disease; cancer; ss

OS Homo sapiens.

PN W0200157271-A2.

PD 09-AUG-2001.

30-JAN-2001; 2001WO-US006662.

PR 04-FEB-2000; 2000US-0180312.
00 MAR 2000 0000US-0207455

PR 30-JUN-2000; 2000US-0608408.
03-JUN-2000; 2000US-0633355

PR 21-SEP-2000; 2000HS-0234687..
DE 27-SEP-2000; 2000HS-02346359

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX

Penn SG, Hanzel DK, Chen W, Rank DK, XY

WPL; 2001-496933/54.

PT New spatially-addressable set of single exon nuclear acid probes, useful for measuring gene expression in sample derived from human

Pt breast, comprises number or single exon nucleic acid probes

PS CLAIM #; SEQ ID NO 0434; 32/PP + Sequence Listing; English
XX

CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived

the probes with a collection of detectably labeled nucleic acids

derived from human breast, and then measuring the label bound to each probe of the microarray. The probes are useful for

CC Verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for

CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 1785 BP; 293 A; 622 C; 480 G; 390 T; 0 other;

Query Match 2.5%; Score 47; DB 22; Length 1785;

Best Local Similarity 49.4%; Pred. No. 0.089;

Matches 122; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 1062 CTTCGGCTTCAGAGCAAGACCTCGAGATAGAGCGCTGCTACCTCAGCGCGAGGTAA 1121
DB 510 CATGCCCTTGAATTGGCCCTACCTGAGAGCTCTTGGCCAGACCTTCATCACTGCTGT 569
QY 1122 AGAGGAG 1181
DB 570 CGGCCACCTGTGCTGCTCAACGACGACGCTCGCTCTATGTCTACCTGCTCTA 629
QY 1182 CCGCTTCGTACGAGATGAGCTGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1241
DB 630 TCTCTTCTTCGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 689
QY 1242 CCGCGCGGAG 1301
DB 690 CTACCTGTGTGCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 749
QY 1302 CTTCGAG 1308
DB 750 CCTGGGG 756

RESULT 15

ABA65633 standard; DNA; 1785 BP.

XX ABA65633;

XX 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #13938.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX Homo sapiens.

XX MO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001MO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human fetal liver -
XX Claim 4; SEQ ID NO 13938; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human fetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a single exon nucleic acid
XX probe of the invention.

XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 1785 BP; 293 A; 622 C; 480 G; 390 T; 0 other;

Query Match 2.5%; Score 47; DB 22; Length 1785;

Best Local Similarity 49.4%; Pred. No. 0.089;

Matches 122; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 1062 CTTCGGCTTCAGAGCAAGACCTCGAGATAGAGCGCTGCTACCTCAGCGCGAGGTAA 1121
DB 510 CATGCCCTTGAATTGGCCCTACCTGAGAGCTCTTGGCCAGACCTTCATCACTGCTGT 569
QY 1122 AGAGGAG 1181
DB 570 CGGCCACCTGTGCTGCTCAACGACGACGCTCGCTCTATGTCTACCTGCTCTA 629
QY 1182 CCGCTTCGTACGAGATGAGCTGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1241
DB 630 TCTCTTCTTCGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 689
QY 1242 CCGCGCGGAG 1301
DB 690 CTACCTGTGTGCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 749
QY 1302 CTTCGAG 1308
DB 750 CCTGGGG 756

Search completed: July 15, 2003, 14:01:32
Job time : 436 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2003, 11:46:09 ; Search time 4928 Seconds

(without alignments)
11037.570 Million cell updates/sec

Title: US-09-966-803-1

Perfect score: 1869

Sequence: 1 ATGACCGCATCGAATGGA...TCAGGGGAATGGAACTGA 1869

Scoring table: IDENTITY_NUC

Gap: 10.0, Gape: 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

- 1: gb_da:*
- 2: gb_hgt:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_da:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_ov:*
- 23: em_ov:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_hgt_hum:*
- 31: em_hgt_inv:*
- 32: em_hgt_other:*
- 33: em_hgt_mus:*
- 34: em_hgt_pin:*
- 35: em_hgt_rid:*
- 36: em_hgt_mam:*
- 37: em_hgt_vrt:*
- 38: em_sy:*
- 39: em_hgt_hum:*
- 40: em_hgt_mus:*
- 41: em_hgtgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1869	100.0	1869	6	AR086952 Sequence
2	1869	100.0	1869	6	AR095647 Sequence
3	1869	100.0	1869	6	AR136115 Sequence
4	1078.2	57.7	304290	1	CNSPAX05
5	1078.2	57.7	349980	6	AX041921
6	955	51.1	257000	1	AP000002
7	825.2	44.2	13354	1	AP010156
8	522	27.9	11370	1	AE013111
9	204.8	11.0	247695	1	AP000064
10	166	8.9	10256	1	AE001879
11	150.8	8.1	265118	1	CNSPAX06
12	150.8	8.1	265118	6	AX041922
13	150.2	8.0	217420	1	BSUB0017
14	149.8	8.0	1971	6	AX006271
15	140.8	7.5	1958	6	AX031809
16	132.2	7.1	11233	1	AE009798
17	131	7.0	251700	1	AP000062
18	120.4	6.4	1896	6	AR116209
19	120.4	6.4	1896	6	E16634
20	120.4	6.4	233000	1	AP000003
21	102.8	5.5	12882	1	AE013176
22	98.6	5.3	9825	1	AE005114
23	96.8	5.2	139203	2	AC068311
24	92.6	5.0	11388	1	AE011750
25	84.6	4.5	2642	1	AF508804
26	79.8	4.3	12098	1	AE012209
27	72	3.9	182440	1	AP000994
28	72	3.9	316900	1	TACID3
29	68.4	3.7	17112	1	AE004997
30	66.2	3.5	477	6	AX436451
31	61.8	3.3	3808	1	AF068743
32	58.6	3.1	12358	1	AE006863
33	58.4	3.1	1836	6	AX078507
34	57.8	3.1	7459	1	SC8A2
35	57.8	3.1	13160	1	SC9C7
36	56.2	3.0	125020	9	AF429315
37	55	2.9	11105	1	AE004963
38	54.4	2.9	39446	1	SCER7
39	53.4	2.9	38634	1	SC26G5
40	52.8	2.8	3217	1	SGSECA
41	52.2	2.8	12212	1	AE005872
42	51.4	2.8	268050	1	AP000983
43	50.6	2.7	26423	1	SC8G12
44	50.4	2.7	125020	9	AF429315
45	50.4	2.7	167446	2	OSJN0085

ALIGNMENTS

RESULT 1
LOCUS AR086952
DEFINITION Sequence 1 from patent US 5985646.
ACCESSION AR086952
VERSION AR086952.1 GI:10013718
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (Bases 1 to 1869)
AUTHORS Murphy, D., Reid, J. and Robertson, D.
TITLE Amidase
JOURNAL Patent: US 5985646-A 1 16-NOV-1999;
FEATURES Location/Qualifiers

source 1..1869
BASE COUNT 476 a 486 c 567 g 340 t
ORIGIN

Query Match 100.0%; Score 1869; DB 6; Length 1869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGACGGGCAATGATGAAACACGACGACCTTTTCTAAGTTCGCTACCTGGGCGACCG 60
DB 1 ATGACGGGCAATGATGAAACACGACGACCTTTTCTAAGTTCGCTACCTGGGCGACCG 60
QY 61 AGGATACGGGCAATGATGAAACACGACGACCTTTTCTAAGTTCGCTACCTGGGCGACCG 120
DB 61 AGGATACGGGCAATGATGAAACACGACGACCTTTTCTAAGTTCGCTACCTGGGCGACCG 120
QY 121 TAGAGAGACAGGCTGTTGTTGAAGACCTTGAACGGGCTCAAGGCGCTTCATGAGAAC 180
DB 121 TAGAGAGACAGGCTGTTGTTGAAGACCTTGAACGGGCTCAAGGCGCTTCATGAGAAC 180
QY 181 GCGCTCAATGCCGAGGATTTGCCAGACGGCAGAAAGCTCGCTTCACCTGCTTTAACGAG 240
DB 181 GCGCTCAATGCCGAGGATTTGCCAGACGGCAGAAAGCTCGCTTCACCTGCTTTAACGAG 240
QY 241 GAGAGAGAGAGACGAGATATGGGTGGCGGATTCAGACCCGAGCGGCAAGAAAGTC 300
DB 241 GAGAGAGAGAGACGAGATATGGGTGGCGGATTCAGACCCGAGCGGCAAGAAAGTC 300
QY 301 CTCTCAACTAAAAAGTCCGCTCGATGATGAGACGATTCAGAGACGATTCATGAT 360
DB 301 CTCTCAACTAAAAAGTCCGCTCGATGATGAGACGATTCAGAGACGATTCATGAT 360
QY 361 GTGGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 GTGGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 TTGCACATATATGGGATTTCTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 TTGCACATATATGGGATTTCTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 GCGGAGAGAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 GCGGAGAGAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 GATGCGATATGTTGGAAGCTCCGCGACGCGGAGGAGGAGCAAGCTCGCTTAAAGTC 600
DB 541 GATGCGATATGTTGGAAGCTCCGCGACGCGGAGGAGGAGCAAGCTCGCTTAAAGTC 600
QY 601 TAGACATATAGTCTATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 TAGACATATAGTCTATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 GAGGCGGTGACTCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 GAGGCGGTGACTCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 ATGACGAGAGACGAGCTGCTGTACTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
DB 721 ATGACGAGAGACGAGCTGCTGTACTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 781 CCGCTCGACGCTCTGGAAGCCAGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 781 CCGCTCGACGCTCTGGAAGCCAGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
QY 841 GCGGAGAGAGGATTAACCTCTGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 841 GCGGAGAGAGGATTAACCTCTGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 901 CACTGATATACGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 901 CACTGATATACGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
```

```
QY 961 ACGAGATATAGGCGAGCTCTACCTCTACAGCGGCGAGCTGAAACAGAGTCCACGATACAC 1020
DB 961 ACGAGATATAGGCGAGCTCTACCTCTACAGCGGCGAGCTGAAACAGAGTCCACGATACAC 1020
QY 1021 GGGCGGATATATCAGGAAGCTCAAGACCTTCAGGCGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1021 GGGCGGATATATCAGGAAGCTCAAGACCTTCAGGCGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 GACCTCGAGATAGACGCGCTGTACCTCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
DB 1081 GACCTCGAGATAGACGCGCTGTACCTCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
QY 1141 ATAGCTTCGTCACGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 1141 ATAGCTTCGTCACGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
QY 1201 CAGCTGATGCGGAGCAAGGCGCTACTACTGCTGCTGGAACCCGCGCGAGGAGGAGGAGGAG 1260
DB 1201 CAGCTGATGCGGAGCAAGGCGCTACTACTGCTGCTGGAACCCGCGCGAGGAGGAGGAGGAG 1260
QY 1261 TATAGCGAAGACTTCGCGCTCCGCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 TATAGCGAAGACTTCGCGCTCCGCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 ATATGAAAGCGCATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
DB 1321 ATATGAAAGCGCATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
QY 1381 GGAATTAACGGGCAATAGCTACGAGGCGCTTCATGACCAACATGAGGCTTACTACAGAGGAC 1440
DB 1381 GGAATTAACGGGCAATAGCTACGAGGCGCTTCATGACCAACATGAGGCTTACTACAGAGGAC 1440
QY 1441 CTCTTCAAGGAGAGATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 1441 CTCTTCAAGGAGAGATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 TCAGACATATAGGCTCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB 1501 TCAGACATATAGGCTCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1561 AACTTCAGGAAGCTACAGCCCGCTGCTTACGCTCAGAAACGTAAGGCGGAGTACTCTTA 1620
DB 1561 AACTTCAGGAAGCTACAGCCCGCTGCTTACGCTCAGAAACGTAAGGCGGAGTACTCTTA 1620
QY 1621 ATCCACGCTTGAAGAGCTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
DB 1621 ATCCACGCTTGAAGAGCTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
QY 1681 CTCAAGACATGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
DB 1681 CTCAAGACATGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
QY 1741 AGCGTCCGCGGAGAGCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
DB 1741 AGCGTCCGCGGAGAGCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
QY 1801 GAGGCAAGCTCAAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
DB 1801 GAGGCAAGCTCAAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
QY 1861 GGAAGCTGA 1869
DB 1861 GGAAGCTGA 1869
```

RESULT 2
AR095647
LOCUS AR095647 1869 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 1 from patent US 6004796.
ACCESSION AR095647
VERSION AR095647.1 GI:10023708
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1869)
AUTHORS Murphy, D., Reid, J. and Robertson, D.
TITLE Amidase
JOURNAL Patent: US 6004796-A 1 21-DEC-1999;
FEATURES Location/Qualifiers
SOURCE 1. 1869
BASE COUNT 476 a 486 c 567 g 340 t
ORIGIN

Query Match 100.0%; Score 1869; DB 6; Length 1869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGACCGGCATCGATGAGACACGACCTTTTCTAGTTCGCTTACCTGGGCGACCG 60
1 ATGACCGGCATCGATGAGACACGACCTTTTCTAGTTCGCTTACCTGGGCGACCG 60
61 AGATACGGGAACTTATCGCGTACACCTGACGAAGGCCAATGAGACAACAG 120
61 AGATACGGGAACTTATCGCGTACACCTGACGAAGGCCAATGAGACAACAG 120
61 AGATACGGGAACTTATCGCGTACACCTGACGAAGGCCAATGAGACAACAG 120
121 TACGAGACGCGGTTGTTGTAAGACCTTGAACGGGCTCAAGGCGCTTATGAGAAC 180
121 TACGAGACGCGGTTGTTGTAAGACCTTGAACGGGCTCAAGGCGCTTATGAGAAC 180
121 TACGAGACGCGGTTGTTGTAAGACCTTGAACGGGCTCAAGGCGCTTATGAGAAC 180
181 GCGTCAATGCCGAGATTTCCGACAGCGCAAGAACTCGGCTTACCTGTTTAAACAG 240
181 GCGTCAATGCCGAGATTTCCGACAGCGCAAGAACTCGGCTTACCTGTTTAAACAG 240
181 GCGTCAATGCCGAGATTTCCGACAGCGCAAGAACTCGGCTTACCTGTTTAAACAG 240
241 GAGAAGAAAGAGACCGGATATGCGGTGGCGCATATCCAGACCTGAGCGCAAGAACTC 300
241 GAGAAGAAAGAGACCGGATATGCGGTGGCGCATATCCAGACCTGAGCGCAAGAACTC 300
241 GAGAAGAAAGAGACCGGATATGCGGTGGCGCATATCCAGACCTGAGCGCAAGAACTC 300
301 CTCTCACTAATAAAAGTCCGCTGATGATGAGAAAGCAAGATTCAGAGACTCTTAACT 360
301 CTCTCACTAATAAAAGTCCGCTGATGATGAGAAAGCAAGATTCAGAGACTCTTAACT 360
301 CTCTCACTAATAAAAGTCCGCTGATGATGAGAAAGCAAGATTCAGAGACTCTTAACT 360
361 GTGCGCTTCAAGAGAGAGAGATGAGACTCTCTTTTGAAGAGAGAGAGAGAGAGAG 420
361 GTGCGCTTCAAGAGAGAGAGATGAGACTCTCTTTTGAAGAGAGAGAGAGAGAGAG 420
361 GTGCGCTTCAAGAGAGAGAGATGAGACTCTCTTTTGAAGAGAGAGAGAGAGAGAG 420
421 TTGACAAATATGGGATTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
421 TTGACAAATATGGGATTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
421 TTGACAAATATGGGATTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
481 GCCGAGAGATATATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
481 GCCGAGAGATATATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
481 GCCGAGAGATATATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
541 GATGCGATATGTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
541 GATGCGATATGTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
541 GATGCGATATGTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
601 TAGGACATATGTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
601 TAGGACATATGTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
601 TAGGACATATGTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
661 GAGGCGGTTACTCCAG 720
661 GAGGCGGTTACTCCAG 720
661 GAGGCGGTTACTCCAG 720
721 ATGACGAGAGACGAGTGGTGTACCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
721 ATGACGAGAGACGAGTGGTGTACCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
721 ATGACGAGAGACGAGTGGTGTACCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
781 CCGCTGAGAGTGGGAG 840
781 CCGCTGAGAGTGGGAG 840
841 GCGGCGAGAGTAAACCTCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900

|||||
841 GCGGCGAGAGTAAACCTCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
901 CACTGAGATTTAGGGCTTACGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
901 CACTGAGATTTAGGGCTTACGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
901 CACTGAGATTTAGGGCTTACGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
961 ACAGAGATATGAG 1020
961 ACAGAGATATGAG 1020
1021 GGGCGGATATTCAG 1080
1021 GGGCGGATATTCAG 1080
1081 GAGCTGAGATATGAG 1140
1081 GAGCTGAGATATGAG 1140
1141 ATAGCTTCTGTCACAG 1200
1141 ATAGCTTCTGTCACAG 1200
1201 CAGCTGATGAG 1260
1201 CAGCTGATGAG 1260
1261 TATAGCAG 1320
1261 TATAGCAG 1320
1321 ATATGAG 1380
1321 ATATGAG 1380
1381 GGAATTAAG 1440
1381 GGAATTAAG 1440
1441 CTCTTCAAG 1500
1441 CTCTTCAAG 1500
1501 TCGGACATATGAG 1560
1501 TCGGACATATGAG 1560
1561 AACTTCAAG 1620
1561 AACTTCAAG 1620
1621 ATCCAGTCTGAG 1680
1621 ATCCAGTCTGAG 1680
1681 CTCAAGGACATATGAG 1740
1681 CTCAAGGACATATGAG 1740
1741 AGGTCGCGGAG 1800
1741 AGGTCGCGGAG 1800
1801 GAGGCGAGAGTCAAG 1860
1801 GAGGCGAGAGTCAAG 1860
1861 GGAAGACTGA 1869
1861 GGAAGACTGA 1869

RESULT 3

ARI36115 ARI36115 .1869 bp DNA linear PAT 16-JUN-2001
LOCUS Sequence 1 from patent US 6136583.
ACCESSION ARI36115
VERSION ARI36115.1 GI:14476787
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1869)
AUTHORS Murphy, D., Reid, J. and Robertson, D.
TITLE AMIDASE
JOURNAL Patent: US 6136583-A 1 24-OCT-2000;
FEATURES
source Location/Qualifiers
BASE COUNT 476 a 486 c 567 g 340 t
ORIGIN
Query Match 100.0%; Score 1869; DB 6; Length 1869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACGGGCAATCGAATGGAACCAAGACACCTTTCTAAGTTGCGCTACCTGGGGGACCCG 60
DB 1 ATGACGGGCAATCGAATGGAACCAAGACACCTTTCTAAGTTGCGCTACCTGGGGGACCCG 60
QY 61 AGGATACGGGGAAAATTATCGCGCTACACCTGACGAAAGGCAATGAAGACACAG 120
DB 61 AGGATACGGGGAAAATTATCGCGCTACACCTGACGAAAGGCAATGAAGACACAG 120
QY 121 TACAGAGACGCGTGTGTGTGAAGACCTTGAACGGGCTCAAGCCCTCATCGAGAC 180
DB 121 TACAGAGACGCGTGTGTGTGAAGACCTTGAACGGGCTCAAGCCCTCATCGAGAC 180
QY 181 GCCCTAATGCCGAGGATTTGCCAGACGGCAAGATCGCTTCACTGCTTTAAGAG 240
DB 181 GCCCTAATGCCGAGGATTTGCCAGACGGCAAGATCGCTTCACTGCTTTAAGAG 240
QY 241 GAGAAAGAGAGACCGAGATTTGGGTGGCGATATCCAGACCTTGAGCGCCAAAGATC 300
DB 241 GAGAAAGAGAGACCGAGATTTGGGTGGCGATATCCAGACCTTGAGCGCCAAAGATC 300
QY 301 CTCTCACTAAAAAGTCGCTCGATGAGTGAAGCAAGCATTCAGAGACTCTTAACT 360
DB 301 CTCTCACTAAAAAGTCGCTCGATGAGTGAAGCAAGCATTCAGAGACTCTTAACT 360
QY 361 GTCGGCTTCAAGAGAGAGAGATGAGACTCTCTTTGACGACGATCCGCTGG 420
DB 361 GTCGGCTTCAAGAGAGAGAGATGAGACTCTCTTTGACGACGATCCGCTGG 420
QY 421 TTGCAACAATATGGAATCTTTGATGAGAGAAAGACGCTTCTGGTCTTGAACAG 480
DB 421 TTGCAACAATATGGAATCTTTGATGAGAGAAAGACGCTTCTGGTCTTGAACAG 480
QY 481 GCCGAGAGAGATTAATCGAGAGTTCGAGAACCGAGTTTTCAGTGGCTTCGACAGC 540
DB 481 GCCGAGAGAGATTAATCGAGAGTTCGAGAACCGAGTTTTCAGTGGCTTCGACAGC 540
QY 541 GATCGATAGTGTGGAAGCTCCGACCGGAGGGGAGCAAGCCTGCTTCAAGTTC 600
DB 541 GATCGATAGTGTGGAAGCTCCGACCGGAGGGGAGCAAGCCTGCTTCAAGTTC 600
QY 601 TACGACATAGTCATATGGAAGAGAGGAGGAGAAAGATCTTCGAGAGGCTCTTC 660
DB 601 TACGACATAGTCATATGGAAGAGAGGAGGAGAAAGATCTTCGAGAGGCTCTTC 660
QY 661 GAGCGGTTGACTCGAGGAAAGAGATTAATCTCTGAGGGGCAAGAAAAAAGCGGTT 720
DB 661 GAGCGGTTGACTCGAGGAAAGAGATTAATCTCTGAGGGGCAAGAAAAAAGCGGTT 720
QY 721 ATCAGCAGACGACTGCTGATCTCTGGAGCGGAGCTTAAACGATCTACGAGGC 780
DB 721 ATCAGCAGACGACTGCTGATCTCTGGAGCGGAGCTTAAACGATCTACGAGGC 780

DB 721 ATCAGCAGACGACTGCTGATCTCTGGAGCGGAGCTTAAACGATCTACGAGGC 780
QY 781 CCCTCGACGCTCTGGAGAGCCAGACGAGGAGGAAAGGTTACTTCTCTACCTCAGAT 840
DB 781 CCCTCGACGCTCTGGAGAGCCAGACGAGGAGGAAAGGTTACTTCTCTACCTCAGAT 840
QY 841 GCGGCGAGGCTTAAACCTCTGCTCTGGAGCGGAGGAGCCGAGCTGTGTTACCGGAG 900
DB 841 GCGGCGAGGCTTAAACCTCTGCTCTGGAGCGGAGGAGCCGAGCTGTGTTACCGGAG 900
QY 901 CACTGATTAACGGGCTTGAAGCTGACGAGGATGGCAAGACATTTGCTCATATGACGCC 960
DB 901 CACTGATTAACGGGCTTGAAGCTGACGAGGATGGCAAGACATTTGCTCATATGACGCC 960
QY 961 ACAGGATTAAGGCGAGCTCTACCTCTACAGCGGAGGCTGAACAGGTCAACGATACAC 1020
DB 961 ACAGGATTAAGGCGAGCTCTACCTCTACAGCGGAGGCTGAACAGGTCAACGATACAC 1020
QY 1021 GGGCGGATTTACAGGAAGCTCAAGACCTTGGAGCCGAGGACACTTCCGCTTCAAGACAA 1080
DB 1021 GGGCGGATTTACAGGAAGCTCAAGACCTTGGAGCCGAGGACACTTCCGCTTCAAGACAA 1080
QY 1081 GACCTCGAGATTAAGGCGAGCTCTACCTCTACAGCGGAGGCTTAAAGAGAGAGGCGCG 1140
DB 1081 GACCTCGAGATTAAGGCGAGCTCTACCTCTACAGCGGAGGCTTAAAGAGAGAGGCGCG 1140
QY 1141 ATAGCTCTGCTCAAGCGCGGCGGAGGAGGATGACGACACCGCTTCTACGAGATG 1200
DB 1141 ATAGCTCTGCTCAAGCGCGGCGGAGGAGGATGACGACACCGCTTCTACGAGATG 1200
QY 1201 CAGCTGATTAAGGCGAGGAGGCTTACTGCTGCTGTGAACCGGCGGCGGAGGAGG 1260
DB 1201 CAGCTGATTAAGGCGAGGAGGCTTACTGCTGCTGTGAACCGGCGGCGGAGGAGG 1260
QY 1261 TATAGCAGAGACTTCGCGCTTCGCGCTCTGAGAGAGACTGGCTGAGAGCTTGAAGAC 1320
DB 1261 TATAGCAGAGACTTCGCGCTTCGCGCTCTGAGAGAGACTGGCTGAGAGCTTGAAGAC 1320
QY 1321 ATATGAACGAGCATACGAGAGTCTTCAAGCTCAACCGGAGGCGGAGGAGGCGGT 1380
DB 1321 ATATGAACGAGCATACGAGAGTCTTCAAGCTCAACCGGAGGCGGAGGAGGCGGT 1380
QY 1381 GGAATTAAGGCGCATACGAGGCTTCATGACCAACTGGGCTTACTACAGAGGAC 1440
DB 1381 GGAATTAAGGCGCATACGAGGCTTCATGACCAACTGGGCTTACTACAGAGGAC 1440
QY 1441 CTCTTCAAGGCGAGATTAAGGAGAGCGCATTAAGTCTGCTACACGCTACGCTTC 1500
DB 1441 CTCTTCAAGGCGAGATTAAGGAGAGCGCATTAAGTCTGCTACACGCTACGCTTC 1500
QY 1501 TCGGACATTAAGGCTCTGCTTACGAGTCAATCGGCGCAATTCGTTAAGAGAGAG 1560
DB 1501 TCGGACATTAAGGCTCTGCTTACGAGTCAATCGGCGCAATTCGTTAAGAGAGAG 1560
QY 1561 AACTTCAGAGAGCTCAAGCCGCTTCTACGCTCAGAACTGAAGGCGCGATACCTCTA 1620
DB 1561 AACTTCAGAGAGCTCAAGCCGCTTCTACGCTCAGAACTGAAGGCGCGATACCTCTA 1620
QY 1621 ATCCATCTGCTTGAAGAGTACCGCTGCTGCTGACAGAGCTTATGTTCTAAAGTG 1680
DB 1621 ATCCATCTGCTTGAAGAGTACCGCTGCTGCTGACAGAGCTTATGTTCTAAAGTG 1680
QY 1681 CTCAAGGACATGAGGAGAGAGAGCTTACATGATCAAGATTTCAAGCGCGGCGCAC 1740
DB 1681 CTCAAGGACATGAGGAGAGAGAGCTTACATGATCAAGATTTCAAGCGCGGCGCAC 1740
QY 1741 AGCGTCGCGGAGAGCCGAGGACAGGCGGAGGCTTACAGGCTTCTATAGATTCCT 1800
DB 1741 AGCGTCGCGGAGAGCCGAGGACAGGCGGAGGCTTACAGGCTTCTATAGATTCCT 1800
QY 1801 GAGCGCAAGCTCAAGAGATTAAGAGAGGCTTATAGATTAAGAGAGGAGAT 1860
DB 1801 GAGCGCAAGCTCAAGAGATTAAGAGAGGCTTATAGATTAAGAGAGGAGAT 1860

/db.xref="SPTREMBL:Q9UZA6"
 /translation="MREPIRTPKGSPRVIFVIGKGVGKTSASISVALAKAGRTII
 VSLDPAHLGVFMERKLSDRPREIINDINVASELDMEGIKRYELHLENKLNMYRYTL
 VINEKEYEYRISGIEBEVATLPAIRILTKGQEMWDYVDPPTGVTMLNMYRYTL
 SLVWEKTEYDTRKILKERAIITLQSGKQVIGGEIILPTREEDAVMELAKAYSD
 EIKFVEDLTPNPKTSYAVANPEPIPLYETKRAYESLKKRIFFNVMYMKVILKLR
 EIPERVLAKQEKYLSKYRKEFPQGVETVELPMQDEPRGIENLSKYGEMLYGR"
 complement(8489..10267).

```

/transl_table=11
/product="CARBON STARVATION PROTEIN A."
/protein_id="CAB50154.1"
/db_xref="GI:5458667"
/db_ref="SPRMBL:09U2A5"
/translation="MNSTVVTLLAAIIITYGLITYFGKGLQKVKVRADPNRTPAHRLX
DGVDPVAPHPVLVYGHHPASTAGAGPIVGPVPAVNAAMGSLPLVNWGVNVEIGAVHDY
IALMSVYDCKSLQWIMAGKLMSKRTSNAEPFLVILITIAFPAALVITFRTGQVAVL
AATSLIELLLAVLLGLMGLIKVQINFKVGTIGILLITIAIITYEKFPLVNAANVTY
FLGAVIYVASSLPVWVLLQDPBDYVLAITMGGVLLIATPIIFGKFGPPTAAVYTT
FSARVAGKRPSPFMTPIPLVLAACGALSGFHSVYSGISSKQDNDELHGLVYCGMFTT
EGFLSTVIITSSIVAYGLQIFNAGIGITAEVMAOAVAPLVNAKSIITTEYLOKLVAAAL
ENFKNGVITGVGVIPASSYANGHLDAPFIDILHTLTPSLISVAFAPALTETDARLQRL
FAMQDIIEAWKGPBELKKNKVAASFLVEFGFGLAMKGGWLVTPAFSGMNMQLSIAM
MTASIVAKTIQRPSCGEMKWAVALPALPPLMVTYVSALIMELIYVCKSLWVSLIITVGL
VLTNELVLDWVNAKKPRAEETQAAAA"

```

```

/transl_table=11
/product-"RNASE L INHIBITOR"
/protein_id-"CAB50135.1"
/db_xref="gi:5458668"
/db_xref="SPTREMBL:Q9U244"
/translation-"MVRKRLAVIYDIDDCNDPKCGHFLCEKRCVCPVNRMGSGAIIIDDE
MKPIIDASCTGGCICVHKCPENASIVYNLPDLDCEINSDMYIAFGNENLONTFELKNG
GMVGVIVPNGTGKTTAVKIIAGLIPMLCNDNSMDVYIAFGNENLONTFELKNG
EIRPVRKQYVDLPEKAVKGVRELLKVDVSGVFAEYVKLEENLVLDRELHDSGG
ELRQVIAAALKRKAHFEDEPSPSYLDIRQLVAVIRLAEAGKVAIVLVEHDLAM
LDYSDVITHVYGEQGVYGFSPKDGKTGNGINEFQGLKIDENYRFRPELRFKLSSE
RDVRELVLEVPELVKVGSPFKLEVEGEIKRGCVGIVGPNIGIKGTPEVKMLAGEV
EPTPECKVSMDLIVAKAKPOYIKAEVBYGYELLSIDSKLSNRYKTELLKPLGIIID
YDRNVEDLSGGEICQAVIAATLLRDADITYLLDEPSALIVDQRLAVSAIHLHKKENK
KTALVEHVDVIMIDVSDRLIVFGSEPEGRHRLAPPMKMRGKNNRFLASVGIETFRRDE
DSGRPRANKEGSVKDEQKARGEAYYA"
13130..13513

```

```

/note="PAB0825"
/codon_start=1
/transl_table=11
/product="TRANSLATION INITIATION INHIBITOR, PUTATIVE."
/protein_id="CAB50156.1"
/db_xref="GI:5458669"
/db_xref="SPTREMBL:Q90ZA3"
/translition="MAKEVIFTEKAPKEIGYSOAIKGNFI'FVAGQIPIDPETGEIYI
RGDIEQRNRVLENIKAILEEGASILNDVVAVTYVLKDLNDFAKMNEYSSYFGESKSK
ARVAAYEVRLPEDVDIEMEAIAIKE"
13623. 14480
/note="PAB0826"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB50157.1"

```

Query Match	57.7%;	Score 1078.2;	DB 1;	Length 304290;
Best Local Similarity	74.9%;	Pred. No. 2.6e-201;		
Matches 1405; Conservative	0;	Mismatches 458;	Indels 12;	Gaps 4;

QY 1 ATGACCGGATGGATGGAACCAAGAGACCTTTCTAAGTCGGCTACCGGGGACCCG 60
210425 ATGACCTCATGGAAGGATGAGAGAGCTTTCAAAAGTTCGGCTCTACTAAGGACCCC 210366
QY 61 AGGATACGGGGAACTTATCGCCCTACACCTGACGAGAGGCCAATGGAAGGACAAAG 120

Db	210365	AGGATTAAGGGGAGACAAATAGCGTAGCTTTCTTAACCAAGGGGAACTTTGACAAACAACAAG	210306
QY	121	TACGAGACGACGCTGTTGTTGTTAGAACCTTTGAACGGGCTCAAGGCGCTTCAATCGAAGAC	180
Db	210305	TACGAGAGCACGATATGTTCTTCGAGAAACTTCGAAAGATGGGAGCCAAAGAGTTCTCATAGAAC	210246
QY	181	GCCCAATAGCCGAGGATTTGCGCAACAGGCAAGAAAGTCGCGCTTCAACCTGCTTTAAAGAG	240
Db	210245	GCTTCATGCTTAGATTTCTTCACAGATGGAAAGAAAGATAGCTTTATATAGAGTTCAACGAG	210186
QY	241	GAGAAAGAGAGACCCGAGATATGGGTGGCCGATATTCAGACCCCTGAGCGCCAGAAAGATC	300
Db	210185	GAGAAAGAAAGTTGCTCAATATATAGGATGGGTCCGGGACATGAAGACCCCTAAGCGCGAAAGATT	210126
QY	301	CTCCTCACTAAACATCCCGCTCCATATGAGTAGTGAAGCAAGCATTTCAAGAGACTCTAGTT	360
Db	210125	CTCAGGCGCAAGAAATATAGGCTCCCTGGAAATGGAATGACACTCAAGAGAGTTGTATGATT	210066
QY	361	GTCGCGCTCAAGAGAGGAGGAGCATAGAGACTTCGTCTTTGACAGACAGCTCCCGGTCTG	420
Db	210065	ATAGGCTTCAGAGAGGAGGAGACATTAAGATTTTCAATATTCAGAGATAGCACTTCACGATTG	210006
QY	421	TTCCACATATAGGGAATCTCTTGATATGAGAGAGAGACAGCTTCTGGTTCTTACACTGAG	480
Db	210005	TTTCGACAACTAGGCGTCTTCTTGACGCGCGAGAAACGACTTCTGATATGATATATCTGAG	209946
QY	481	GCCGAGAGATATATGACAGCTTCGAGAAAGCCGAGATTTTCAGTAGCCCTCTGCGACGCG	540
Db	209945	GCCGAGAGGTTATATGACAGAGTTGAGAAAGCCAAAGTTTTCGAGGCGGATTTGGCACGGG	209886
QY	541	GATCGGATATGTTGTATACAGTCCCGCACCGCAGGGGAGCAACCTGCCCTGTCAAGTTC	600
Db	209885	GATTCGATATGTTGTATACGTCCTCCGTACAG--GGACACCATACCAAGGTACTTCAAGTAC	209829
QY	601	TACGACATATGCTTATGAGAGAGCGGGAGAGAAAGAACTCTCGAGAGGCTCTCTTC	660
Db	209828	TGGAACATCTACCTCTGGAAGAGATGGAGAGGAAGAAAGCTATTCGAGAAGGTTCTTTC	209769
QY	661	GAGGCGGTTGACTCCGACGGAAGAAATATCTCTGAGGGGCAAGAAAAAAGCGGTTTC	720
Db	209768	CACGCTATAGACTCTGACGAGGAAGAAATATCTCACTCTACGGAAGGCCCGAGAAAGAGTAC	209709
QY	721	ATACGCGAGCAGACTGGCTGTACCTACCTCGGAGCGGAGACTTAA---ACCGATCTACGAG	777
Db	209708	ATGACCGAGCAGATTAAGCTCTACCTCTACCTACGAGCGGAGAGGTTAAGGAATTAATGATTGC	209649
QY	778	GGCCCGCTGACGCTGTGGGAAGCCAAAGCTCACAGGAAGAAAGTCTACTCTCACTCCA	837
Db	209648	ATAGACAGAGGAAGCTGTGTCACACTTAAGATCAAGGATGGGAAGGTGATCTTCACTCACTAC	209589
QY	838	GATGCGGGCAGGCTAAACCTCTGCGCTCTGGAACGGGAAGGCCGAGCGGTGTGTACCGGC	897
Db	209588	GAGGAGGGAAGGTTTAACCTATATCTCTGGGATGGGGAATATCAAGGATATACTTAAGGCG	209529
QY	898	GACCACTGATTTACGGGCTTGACCTCAAGCTCAAGCGATGGCAAAAGCATTTGCTCTCATCATGACC	957
Db	209528	AAGCACTGATTAATGGGGTTTCACG---TGACGAAGAAAGTTGTCTACTCTCAAGGAAACT	209472
QY	958	GCCACGAGAGATAGCGAGCTCTACTCTATACGACGCGGACACTGAACACAGTCAACCGAATAC	1017
Db	209471	GCCACTAGGTTGAGGAGAGCTCTACATCTGGGACGGGAGAGAGGCAACTATACGGATTAC	209412
QY	1018	AACGGGCGGATTTATAGGAAGCTCAAGACCTTCGAGCCGAGGCACTTCGCTTCTCAAGAC	1077
Db	209411	AACGGCTCGATTTATAGCAAGCTCAAAACCTTCGAGCCCTAAGCACTTCAAGTTCAAGAGG	209322
QY	1078	AAAGACTCGAGATAGACGGCTGTACTCTCAGGCGGAGGTTTAAAG---GAGAAGGCC	1134
Db	209351	ATCGACTTGGAGATAGACGGCTGTGTATCATTAAGGCCCGAAGCTCAAGAGGCGCAAGAGGCT	209292
QY	1135	CCGGGATATAGCTTGTGTCACGGGGGGGCGGAAGGGGATATACGGACACCGCTTCGCTAC	1194
Db	209291	CCGGTAATAGCTTGTGTCACGGGTGGGCCCAAGGGGATATAGCGTTACTACTTCAAGTAC	209232

Db	205348	GAGGAGGGGAACGGTTAAACCTTATCTCTGGGATGGGGATATACAGAGATACCTAAGGCG	205289
OY	898	GACACATGCATTTTACGGGCTTGACGCTCAGCGATGGCCAAAGCATTTGCTCTCATCATGAC	957
Db	205288	AAGCACTGTGAATATGGGGTTGACGT---TGACGGAAGGTTGTCTACTCTCAAGAAACT	205232
OY	958	GCCACGAGGATAGGCGACCTCTACTCTACGACGCGCGAGCTMAACAGGCTCACCGAATAC	1017
Db	205231	GCACATAGATTGAGGAGACTCTACATCTGGGACGGCGAGAGAGGAACCTAACCGATTAC	205172
OY	1018	AACGGGCCGATATTTCAGAAAGCTCAAGACCTTCGACCGGAGCACTTCGCTTCANAGAC	1077
Db	205171	AACGGTCTGAATATTCAGCAAGGCTCAAAAACCTTCGAGCCCTAACACACTTCAAGTTCAAGAC	205112
OY	1078	AAAGACCTCGAGATATGACGGCGGTGACCTCAGGCGCGAGAGTTAAAGAG--GGAAGGGC	1134
Db	205111	ATCGACCTTGGAGATATGACGGCGGTGACATTAAGCCCGACCTCAAGGAGGCGGAAAGGCT	205052
OY	1135	CCGGTGAATAGTCTTCTGCTCCACGGCGGGCCGAAAGGGCATGTACGACACCGCTTCTAC	1194
Db	205051	CCGGTAATAGACTTCTGCTCCAGGGTGGGCCCAAGGAAATGTACGGTTACTTCACTTCAAGTAC	204992
OY	1195	GAGATGCAAGCTGATGGGCGAGCAAGGGGCTACTCTGCTGCTGTGTGAACCGCGGGCGAC	1254
Db	204991	GAGATGCAAGCTTATATGGCGAGCAAGGGGCTACTCTCATATGATATGTCAATCCAAAGGAAAG	204932
OY	1255	GACGGCTATTAGCGAAGACTTCGCGCTCGCGCTCTGAGAGGAGTGGCTTGGAGACTT	1314
Db	204931	AACGGTTACACGAGAGACTTTTGCACTTATAGAGTCTCTGAGAGGAGTGGCTTACAGGACTTTC	204872
OY	1315	GAGGACATAATGAAGGGCATGAGAGATTTCTCAAGCTCGAACCAGAGGCCACAGAGAG	1374
Db	204871	CAGGATATATCTGAATGGAATGGAATGAGAGATTCTCAAGCTTGAACCGGAACCGGACAGAGAG	204812
OY	1375	CGCGTTGGAATAAACGGGCAATAAGCTACGGCGGCTTCATGACCAACTGGGCTTGACTCAG	1434
Db	204811	AGGGTTGAATTAACGGGAATTAAGCTACGGAAGGCTTCATGACGAAGTGGGCCCTTAACGACAG	204752
OY	1435	AGCGACCTCTTCAAGCGAGGAATTAAGCGGAAGACGCAATAGCTACTGCGTCAACCAACTAC	1494
Db	204751	AGCGACCTATTTCAAAGCCGGAATACGCGAAGACGGAATTAAGTACTGCGTAAACGAAGTTAC	204692
OY	1495	GCGTTCTTCGACATATAGGGCTCTGTGTACGACGCTCGAAGGTCATGGGGCCAATCCGTTAGAG	1554
Db	204691	GCGTTCTTCGACATATAGGATTTGTGTGTTGACAAAGGAAGTTATATGTTAGAAACCCGCTAGAG	204632
OY	1555	AACGGAACCTTCAGGAAGCTCAGCCCGCTGTCTACGCTCAGAACGTTGAAGGCGCCGATA	1614
Db	204631	AACGGAACCTTCAGGAAGCTTAAGTCCGCTTCTACGCTTAAGAACTGAAGGCTCCCGTG	204572
OY	1615	CTCCAAATCCACTCCCTTGAGAGACTACCGCTTCGCTCGCTGACAGAGCTTATGTTCTAC	1674
Db	204571	TTTGCTAAATCCACAGCCTTGGAAGATTATGCTCTCCATTGGAATCAGAGCTTGATGTTCTAC	204512
OY	1675	AACGTGCTCAAGGACATGGGCGAAGAAAGCTTACATAGCGATATTCAAAGCGCGGGCGCCAC	1734
Db	204511	CACGCTCTGAAGAAGATTATAGCAAAAGTTTACATAGCGATATTCAAAGAGGGGAGGCTCAT	204452
OY	1735	GGCCACACAGCTCCGGGGAAGCCCGGAGGCAACAGCGCAACGCTTACAGGCTCTTCATATAGAG	1794
Db	204451	GGACATACATTAAGGGGAAGCCCAAGACACAGAGATTAAGAGGATNCAAGCTCTTCATATAGAG	204392
OY	1795	TTTCTTCAGAGCGCAAGCTCAAGAAATGACGAAGGAGGGCTTTGAGGTTAGAGAAAGATACTCAAG	1854
Db	204391	TTTCTTCAGAGAGAACTCAAGAAATGACGAAGGAGGGCTTCGACGTTGAGAAAGATATCTTAAG	204332
OY	1855	GGGAATGGGAAGCTGA 1869	
Db	204331	GGTGGAAACCAAGTGA 204317	

AP000002	LOCUS	257000 bp	DNA	linear	BCT 06-APR-2000
DEFINITION	Pyrococcus horikoshii OT3 genomic DNA, 287001-544000 nt. position (277)				
ACCESSION	AP000002	AB009475	AB009476	AB009477	AB009478
VERSION	AB009481	AB009482	AB009483	AB009484	BA000001
KEYWORDS	AP000002.1	GI:3236129			
SOURCE					
ORGANISM	Pyrococcus horikoshii (strain:OT3) DNA.				
REFERENCE	Pyrococcus horikoshii				
AUTHORS	Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae; Pyrococcus.				
TITLE	1 (sites)				
JOURNAL	Kawarabayashi, Y., Sawada, M., Horikawa, H., Halkawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohiku, Y., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushiida, N., Oguchi, A., Aoki, K., Nakamura, Y., Robb, T. F., Horikoshi, K., Mauchi, Y., Shizuya, H. and Kikuchi, H.				
MEDLINE	Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3				
REFERENCE	DNA Res. 5 (2), 55-76 (1998)				
AUTHORS	98344137				
TITLE	2 (bases 1 to 257000)				
JOURNAL	Tanaka, T., Kawarabayashi, Y. and Kikuchi, H.				
MEDLINE	Direct Submission				
AUTHORS	Submitted (11-JUN-1998) Yutaka Kawarabayashi, National Institute of Technology and Evaluation, Biotechnology Center, 2Chome 49-10				
TITLE	Nishihara, Shinya-Ku, Tokyo 151-0066, Japan				
JOURNAL	(E-mail: genomeOT3@nate.go.jp, Tel: +81-3-3461-8951, Fax: +81-3-3461-8424)				
COMMENT	On or before June 28, 1998 this sequence version replaced				
	gi:3130553, gi:3130577, gi:3130627, gi:3130656, gi:3130685, gi:3130705, gi:3130737, gi:3130795, gi:3130824, gi:3130828.				
	AB009475-AB009484: submitted (10-DEC-1997)				
	Kawarabayashi, Y. is officially affiliated with the National Institute of Bioscience and Human Technology, Tsukuba, Ibaraki 305-0046, Japan.				
	Robb, T. F. is at the Center of Marine Biotechnology, University of Maryland, Baltimore, MD, USA.				
	Horikoshi, K. is at the Japan Marine Science and Technology Center, Yokosuka, Kanagawa 237-0061, Japan.				
	Nishihara, Y. is at the University of Tokyo, Meguro, Tokyo 153-0041, Japan.				
	Shizuya, H. is at the California Institute of Technology, Pasadena, CA, USA.				
	The other authors are at the National Institute of Technology and Evaluation, 2-49-10 Nishihara, Shinya, Tokyo 151-0066, Japan.				
	All the sequence with length 100 codons or more between ATG or GTG and stop codon are defined as CDS				
	Homology analysis is performed by Smith-Waterman algorithm against Genbank and Genpept release 103; EMBL release 52.0; Swissprot release 34.0; PIR-Protein release 54.0; and OML release 29.5.				
	E-mail address for comments and questions: genomeOT3@nate.go.jp				
	Restriction map, ORF organization, sequence alignment and more information are available at M.W.M. site of Biotechnology Center, URL: http://www.bio.nite.go.jp/ .				
FEATURES	Location/Qualifiers				
Source	1..257000				
	/organism="Pyrococcus horikoshii"				
	/strain="OT3"				
	/db_xref="taxon:53953"				
	/map="287001-544000 nt. position"				
	1363..2664				
	/gene="PH0318"				
	1363..2664				
	/gene="PH0318"				
	/note="Similar to PIR:E6485 percent identity: 55.181 in 403aa; PIR:S74842 percent identity: 49.869 in 397aa; Swiss_prot:P33221 percent identity: 50.773 in 404aa"				
	/codon_start=1				
	/transl_table=11				
	/product="433aa long hypotheticalal				

phosphoribosylglycinamide formyl transferase"
 /protein_id="BAA29392.1"
 /db_xref="GI:3256709"
 /translation="MVMYIKRLDELGATATDSNOKILLGSGELGRKAIENORLGYE
 VVAVDRYANAPAMOVARSYVNMMDKDFLMSYVEREKPDALIPETIAINDALFEPE
 KDCGYVVPNAPATMIAMHRRERLTLYREKAVPTSRMTATTDDELTECEKIGTPOH
 TKAIMSSSGSYTFVKGPEDI PRAMEAKTKARSAKIIIEHIDPDEVTELAVRH
 FDENGELVTTFPKFVGHYQIDGDYHASMOPAEI SEKAEREVYRIAKRITDVLGIGIF
 GVEAFVGDVWVANEVSPRPHDTGMVLASHPQSEFELRAVLGLPIPIGEVVDY
 RLEPMLIPATHVYIKAKVSGSPRFRGLVLAISVNAVTPLFGKPEAVYGRRLGIALA
 MCKDVEAKRKAEEVAMHIELRTSRSSMDHONTEKRHLIR"
 complement(1645. .1974)
 /gene="PH0319"
 complement(1645. .1974)
 /gene="PH0319"
 /codon_start=1
 /transl_table=11
 /product="109aa long hypothetical protein"
 /protein_id="BAA29393.1"
 /db_xref="GI:3256710"
 /translation="MFLNYILCTTSRPSLCFFPGFNNVLGPLEVRSLSRGARYSLC
 VAMITDLETLIEFIKGCYHIFGSRMLCFLHKGFPRLPVHSPGSSCLMNTKITLL
 LEPEEGI"
 complement(2653. .3186)
 /gene="PH0320"
 complement(2653. .3186)
 /gene="PH0320"
 /note="similar to owl:SSOUREK percent identity: 68.000 in
 150aa; PIR:s76001 percent identity: 60.694 in 173aa;
 owl:BAPURKEI percent identity: 62.329 in 146aa"
 /codon_start=1
 /transl_table=11
 /product="177aa long hypothetical
 phosphoribosylaminoimidazole carboxylase catalytic
 subunit"
 /protein_id="BAA29394.1"
 /db_xref="GI:3256711"
 /translation="WYISMSSEKPIVGIINGSDSDPVYKEARILREGVPEITIT
 SAHRTPERAETAKKAERIEVITAGAGAAHLPGTASLTVLPVIGVPIKSKALNG
 LDSLSLIVQMPGSIPIVATVAIDAKNAALALRLITGIKPEIKELRRYKDKMRKVE
 EKARLSIEQMGWERYLSE"
 complement(3275. .4036)
 /gene="PH0321"
 complement(3275. .4036)
 /gene="PH0321"
 /codon_start=1
 /transl_table=11
 /product="253aa long hypothetical protein"
 /protein_id="BAA29395.1"
 /db_xref="GI:3256712"
 /translation="MISWEIEELDREITGKIKHSLILIHEDASSRQKDLITLRSK
 LKSDNLVGMESISYPLQILIRISRGVDVITKLENNRLIIVTFPSFHGIKATMGCV
 WYLEGMSSETLPIKAKAVEDHKVMDLNLFEGRRLYGFALISMSGYLVEFPEETL
 RYLETSAEYVGHAPAKKYPRGFNFMMEGVKXKRYLLSYRRADYVTKRRSLGENG
 IKRELIVTKPKRIEELVREVEYFKNGEPLKRYVD"
 complement(4039. .4698)
 /gene="PH0322"
 complement(4039. .4698)
 /gene="PH0322"
 /codon_start=1
 /transl_table=11
 /product="219aa long hypothetical protein"
 /protein_id="BAA29396.1"
 /db_xref="GI:3256713"
 /translation="MGIIELVSSIEPGGILAIIOKELESEGDREGLVYLKHLLEN
 NVLEFLYEPLRTFMNMINNIGINVEYLGNLIIFFVFGSMNTEKRIKGVITLISGL
 DVVEYSKLKEMAKITLKTNTINANWFEPFLSSGICISNPLITKTIAMALEETIK
 GYRKRIITYISQECPILEETVYLADYILERTINIGKAVGIITKGPENMLIFEIKRE
 V"
 4856. .6172
 /gene="PH0323"
 4856. .6172
 /gene="PH0323"
 /gene="PH0323"

/note="similar to PIR:A64417 percent identity: 55.478 in
 442aa; owl:HUMRPS4 percent identity: 33.738 in 424aa;
 Swiss_prot:P12039 percent identity: 33.819 in 352aa.
 motif-prokaryotic membrane lipoprotein lipid attachment
 site"
 /codon_start=1
 /transl_table=11
 /product="438aa long hypothetical
 phosphoribosylamide-glycine ligase"
 /protein_id="BAA29397.1"
 /db_xref="GI:3256714"
 /translation="MKVILVGGGGRHAIGFALYKGAELVYVSNHRNPGLMRLANDY
 GLAKETNVEDIVKFAKKGIELAFIGPEALENGIVALEEGIPAGPREARLET
 NKWAREREMERNNIPGRKMRIPDDVOEMRMVTEGKPVVVDIGLITGGKGVKGY
 OKMNEAEKYEAEHITRKDGVLEIEEDVCEVETLQYTFGGKVIIPPLVODYPHAE
 GDVPTITGGKMSYSCSNHLPTFTGSPERLKTLEETIEMAREGPTVGIILGDM
 LSSEGPDLIETNAREGDEAIINVLADNDILAKIVGSLRKAKAEIKARAVVYTI
 APOGYPODPIKGIKRIEVDDEGINKEGAKIYAAVDMLTLAGSALAVGADSLSEA
 ERIENGVSYVKGPIFYRKDVGTRESVEKRIEIKRKKEFEPLIC"
 6174. .7325
 /gene="PH0324"
 6174. .7325
 /gene="PH0324"
 /note="similar to PIR:H64316 percent identity: 44.847 in
 382aa"
 /codon_start=1
 /transl_table=11
 /product="383aa long hypothetical protein"
 /protein_id="BAA29398.1"
 /db_xref="GI:3256715"
 /translation="MYIMIDREELSLTEKYDKITGTVIGSHSALIDAKAREGF
 PVLVAAGRHRYAKYKFKDKNDITKGFIDVILIDFAOIIIEIOEELIRNVI
 IPIRSFVYTGIDRENEFVPMFGTSLRTEREEREKSYMLKRAGLPYEEKP
 EIIDVGLVAVKLPKARKLERGFTTASVSKERKESERILRIGVITRHIDARIER
 YIIQVNPDEPFEYSPIDEETIELGVMPRETSIDGHRLEPAAOULIPHQPEPTTY
 CGHASATLRESLEKYVDMAEKTYEAKTKITPGLIDPFTLQTVVDKDLVLFSDAP
 RIGGGINIHMSKMGHPGNSLMRSPMSGRIALTEIKGILDELEKLYT"
 7392. .8486
 /gene="PH0325"
 7392. .8486
 /gene="PH0325"
 /codon_start=1
 /transl_table=11
 /product="364aa long hypothetical protein"
 /protein_id="BAA29399.1"
 /db_xref="GI:3256716"
 /translation="MSSAKIIEFYDEXSDSYVAELAKYENKEIREPIRLSRM
 EAKHAKWDFLVARGYKPKPRIKTLFTVTKLRLFGEPANVSLKENGSAIQ
 YKPFGEESDEITEEPEGLRGIITIDLEHEKFSSESKLFEHENTRDVFLGNDDLV
 ELIGAVYGLSAVYPNSPRLVIGLIVAGVALSMGIALISVRSOROVSEAIERTK
 ILRFVSEKAKEELYERLVGGLPEELARAVSEKLEKEAIIKLIVPEERNEFPAA
 LYGIALVFGVAPVPVYFPAASVILRALPSVTLAGALAIIVATSSILSGISIRKKV
 IEMVTGIGAAFLSYIFGHIMESLFINVASL"
 8483. .11494
 /gene="PH0326"
 /gene="PH0326"

Query Match 51.1%; Score 955; DB 1; Length 257000;
 Best Local Similarity 70.8%; Pred. No. 3,7e-177;
 Matches 1328; Conservative 0; Mismatches 555; Indels 12; Gaps 4;

QY	1	ATGACGGGCAATCGAATGGAACACGAGACCTTTTCTAAGTTCGCTACCTGGCGGACCCG	60
DB	246034	ATGACCTTCATCGAATGGATGAAACACCTTTACGAAGTTCCTACCTACGACCC	246093
QY	61	AGGATACGGGGAACCTTAATCGGGTACACCTGACGAGGCAACATGAAGGACCAACAG	120
DB	246094	AGGACAAAGGAAACCTTGCTTCCTACCTTCAACAAGGCAACCTGATGACGAATPAG	246153
QY	121	TACGAGACGACGGTGTGTTGTAAGACCTTGAACGCGCTCAAGCGGCTTCATCGAAGAC	180
DB	246154	TATGAGACACACTATAGTTATAGAAACCTTGAAGATGGAGACGAGAGTTCATAGAAAGAT	246213
QY	181	GCCCTCAATGCCGAGGATTTGCCACGACGACGAGAAAGCTGCCCTTACCTGCTTTAAAGAG	240

Db 246214 GCTTCATGCGAAGATATCTCCGATGGAAGAGATAGCCTTCTGATGATTCACGAG 246273
 Oy 241 GAGAGAGAGAGACCCAGATATGGGTGGCCGATATCCAGACCCGACGAGAAAGTC 300
 Db 246274 GAGAGAGAGAGACCCGATATGGGTGGCCGATATCCAGACCCGACGAGAAAGTC 246333
 Oy 301 CTCCTCACTAAAGACCTCCGCTCGATGACAGTGAACAGAGATTCAGAGAGACTCTTAA 360
 Db 246334 CTCGAGGCCAGAAATTAAGATCCATAGAGTGAACAGACTCAAGAGAGGCTTCTCGCT 246393
 Oy 361 GTCGCTTCAG 420
 Db 246394 GTTGGGTTAAAG 246453
 Oy 421 TTCGACAAATATGGGATTTCTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 Db 246454 TTCGACAAATATGGGATTTCTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246513
 Oy 481 GCCGAGAGATATGAG 540
 Db 246514 GGAGAGAGAGATATGAG 246573
 Oy 541 GATGCGATATGATGAG 600
 Db 246574 GATTCATATGATGATGAG 246630
 Oy 601 TACGACATATGATGATGAG 660
 Db 246631 TACGACATATGATGATGAG 246690
 Oy 661 GAGGCGATGATGATGAG 720
 Db 246691 TACGACATATGATGATGAG 246750
 Oy 721 ATCAGGAG 780
 Db 246751 GATGCGATATGATGAG 246810
 Oy 781 CCGCTC--GACGCTGAG 837
 Db 246811 ATGACAG 246870
 Oy 838 GATGCGAG 897
 Db 246871 GAGGAG 246930
 Oy 898 GACCACTGATATGAG 957
 Db 246931 GACCACTGATATGAG 246987
 Oy 958 GCCACAGAGATAG 1017
 Db 246988 GCGACATGAG 247047
 Oy 1018 AACGGCCGATATGAG 1077
 Db 247048 AACGGCCGATATGAG 247107
 Oy 1078 AAAGACCTGAGATAG 1134
 Db 247108 ATGACAG 247167
 Oy 1135 CCGGATATGATGATGAG 1194
 Db 247168 CCGGATATGATGATGAG 247227
 Oy 1195 GAGATGAG 1254
 Db 247228 GAGATGAG 247287
 Oy 1255 GACGCTATGAG 1314

Db 247288 AACGCTACAGAGAGATTTTCCGCTGAGAGCTTACAGAGAGAGAGAGAGAGAGAG 247347
 Oy 1315 GAGACATATAG 1374
 Db 247348 CAGGATATATGAG 247407
 Oy 1375 CCGGCTGATATAG 247467
 Db 247408 AGATGCTGATACAG 247467
 Oy 1435 AGCAGACTTTCAG 1494
 Db 247468 AGTACACTTTCAG 247527
 Oy 1495 GCTTCTGAG 1554
 Db 247528 GCTTCTGAG 247587
 Oy 1555 AACGAGACTTTCAG 1614
 Db 247588 AACGAGACTTTCAG 247647
 Oy 1615 GCTCTATCAG 1674
 Db 247648 CTAAGTATCAG 247707
 Oy 1675 AACGAGACTTTCAG 1734
 Db 247708 CAGGCTCTCAG 247767
 Oy 1735 GGCACAG 1794
 Db 247768 GGCACAG 247827
 Oy 1795 TTCTTTCAG 1854
 Db 247828 TTCTTTCAG 247887
 Oy 1855 GGAATGAG 1869
 Db 247888 GAG 247902

RESULT 7
 AE010156
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Pyrococcus furiosus DSM 3638.
 Pyrococcus furiosus DSM 3638.
 Archaea: Euryarchaeota; Thermococci; Thermococcales;
 Thermococcaceae; Pyrococcus.

REFERENCE
 1 (bases 1 to 13354)
 Maeder, D.L., Weiss, R.B., Dunn, D.M., Cherry, J.L., Gonzalez, J.M.,
 Diruggiero, J., and Robb, F.T.
 Divergence of the hyperthermophilic archaea *Pyrococcus furiosus* and
P. horikoshii inferred from complete genomic sequences
 J. Mol. Biol. 330, 134-157 (2001)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

3 (bases 1 to 13354)
 Weiss, R.B., Dunn, D.M., Robb, F.T. and Brown, J.R.

TITLE	The complete sequence of the <i>Pyrococcus furiosus</i> genome	
	Unpublished	
REFERENCE	4 (bases 1 to 13354)	
AUTHORS	Weiss, R.B.	
TITLE	Direct Submission	
JOURNAL	Submitted (12-FEB-2002) Human Genetics, University of Utah, 20 South 2030 East, Salt Lake City, UT 84112, USA	
FEATURES	Location/Qualifiers	
SOURCE	1. 13354	
gene	/organism="Pyrococcus furiosus DSM 3638"	
	/strain="DSM 3638"	
	/db_xref="taxon:186497"	
CDS	94. 1113	
	/gene="PF0308"	
	94. 1113	
	/gene="PF0308"	
	/note="Function Code: 9.3 Purines, Pyrimidines, Nucleosides and Nucleotides: Purine ribonucleotide biosynthesis: (pura)"	
	/codon_start=1	
	/transl_table=11	
	/product="adenylosuccinate synthetase"	
	/protein_id="AAL80432.1"	
	/db_xref="GI:18892258"	
	/translation="MPSVIYVGGMGDEKSGITATIALHDEPEITARGVGYNAGHS VYINGKRYAVROLPTGPMQKRALIVAGVYVDPPEVFFTELEHDKDFNVAERGVIDIR CAIIEPRHKMKNDLHGISTGTSGGCPANADRYMRAKQAKDIKELEPLTDVA AEVNDLDEGALVLEGTQGFGLSYGYGYVYTSKDTASAVASDVGIGPTVDVI VVKSSEPTRGVAGCPPTGEMERADRLGLIEYGVTVGRRRVGMFDEMARVARIAG ATMTAVTMDKYEKAFGITDYDKLPRKAKEFIEIEERGVGVGLIKTGPLEHHID LRENI"	
tRNA	complement(1174. 1251)	
	/product="tRNA-Glu"	
gene	/note="tRNA Glu anticodon TTC"	
	complement(1303. 2580)	
CDS	/gene="PF0309"	
	complement(1303. 2580)	
	/gene="PF0309"	
	/note="Function Code: 14.6 Transport and Binding Proteins: other"	
	/codon_start=1	
	/transl_table=11	
	/product="ABC transporter (membrane protein)"	
	/protein_id="AAL80433.1"	
	/db_xref="GI:18892259"	
	/translation="MSDFVLTVYKEKLDLRDGLIFGIIIVPLIITPALGQMLQVGF EDAQETKAVLVNDEGRYGNILIKALSNPNTVTKIEAKTIDEALKRAQEEYNNI VIIPKNSQAIENSKASVEIYIGIVSGSMKASSEGNNIVISILNEIYAKTIE RRIEDPEALKRIDAKSYTIKRIYVEIPSVSNIIASQSPSMELVIFIMIMYAO MSAGSMAMERKNTLETLLTLPYKRITIVAGVGAIIIGIIAISYMGKNTYGLK TSSAOYSTLEELGKITPLGATLPIIIIFLVTVLSRAMLAIVAEQDKTANTVY AGIMPLAFTFILMPADVNTLPVLKLLGLIPFSPHILASRAMLEEYSTMVASVY LGILSAVTLPEVTAKEPTERKLTAKLRFMKKE"	
gene	complement(2573. 3442)	
	/gene="PF0310"	
CDS	complement(2573. 3442)	
	/gene="PF0310"	
	/note="Function Code: 15.2 Other Categories: Drug and analog sensitivity"	
	/codon_start=1	
	/transl_table=11	
	/product="daunorubicin resistance ATP-binding protein homolog drrA"	
	/protein_id="AAL80434.1"	
	/db_xref="GI:18892260"	
	/translation="MGERVDFSCGWINRVWRILIKDFRSSTFWSYANNSWCHGEGK MKAIVDELDKDYKALKGISFEVKEGEIFGLGPNAGKSTTLKIATLTLVPPGG RAEIFGYDVEAEYKRLISYLEPAGAYKNTGYELEPMAKIYAKGKKYKKEK NGVELSGEERLKDRISTYSGMTRKLLARALMPEKAIIDEPASGDIINASIR OTIKOPARKGTTPLSSNMLEVEPLCDRYALLINKRIVEITGPTELKYNAMENLE EYFMSVGGKAYE"	
gene	complement(3307. 3657)	
	/gene="PF0311"	
CDS	complement(3307. 3657)	
	/gene="PF0311"	
	/note="Function Code: 16.1 Conserved Hypothetical"	
	/codon_start=1	
	/transl_table=11	
	/product="hypothetical protein"	
	/protein_id="AAL80435.1"	
	/db_xref="GI:18892261"	
	/translation="MMKIRKIGIITLTGFTLVIRPPDYRACMTLPGEYCSGI NMPELTGTVYVAGTLITLGNFRKCKERGNIFLAGGLGIGFICGYSRIELVAF GAMLVTIGVMEVRK"	
gene	complement(3724. 5091)	
	/gene="PF0312"	
CDS	complement(3724. 5091)	
	/gene="PF0312"	
	/note="Function Code: 6.7 Energy Metabolism: Glycolysis"	
	/codon_start=1	
	/transl_table=11	
	/product="Adp-dependent glucokinase"	
	/protein_id="AAL80436.1"	
	/db_xref="GI:18892262"	
	/translation="MPWEEYKRNATEKATKSVKYGVGLTYNTNDAIKYDSHL EERLIKREYIKYSRELDKINTYSQGLSITMSIRKKALELVESCPVFTMR WGNELRMGQAGIMANLIGGVGVPTVAVPOLSRLOMLFDGPIYPTLEGEYK LIHPKESGDENCIHYIYEPFGFRFEAPARENFTISADYNTTLTIRREFRS FSEYIKNVQALISGLQALKEKYEFEELKSNLEVNREIIPVHLEAFYDEKVR EELTNVGMFYSVGLNEVELASIMEIIGERKLAKELAHDPVPDIATVETAMLAKT GKVRHHTGYVIALIYEVYGEVRDLPLPALAAAKAKKGNITSIEIIRATSPVY NERATQVEELRLREYIGKEIGVEGQIALPIFKIYAKKRSVIGIGDTISSAFIGE FSTPL"	
snRNA	complement(5145. 5199)	
	/note="small nucleolar RNA, posttranscriptional modification of rRNA"	
gene	complement(5192. 5449)	
	/gene="PF0313"	
CDS	complement(5192. 5449)	
	/gene="PF0313"	
	/note="Function Code: 4.6 Cellular Processes: Protein and peptide secretion; (sec11)"	
	/codon_start=1	
	/transl_table=11	
	/product="signal sequence peptidase I"	
	/protein_id="AAL80437.1"	
	/db_xref="GI:18892263"	
	/translation="MYENPLYKYPITIHRYVEIKTISIEGHEECFTWGDNNPVPDY PPYGMIDCVPGAIRAKALVVPFRIGIISIKVREFLGIG"	
gene	complement(5504. 5713)	
	/gene="PF0314"	
CDS	complement(5504. 5713)	
	/gene="PF0314"	
	/note="Function Code: 16.1 Conserved Hypothetical"	
	/codon_start=1	
	/transl_table=11	
	/product="hypothetical protein"	
	/protein_id="AAL80438.1"	
	/db_xref="GI:18892264"	
	/translation="MHGNNNSREDKNNNGVYMERAVREILSTVLVPLTSSYYGM RIALHTDPLVAVVIOLVHFECLOR"	
gene	complement(5641. 6186)	
	/gene="PF0315"	
CDS	complement(5641. 6186)	
	/gene="PF0315"	
	/note="Function Code: 16.1 Conserved Hypothetical"	
	/codon_start=1	
	/transl_table=11	
	/product="hypothetical protein"	
	/protein_id="AAL80439.1"	
	/db_xref="GI:18892265"	
	/translation="MTIGLVNFTYDTRKLIHEAHIRAIARAAYAVYERHALLIGFPF DSDDEKVEARVASTTTIGEGKRYLELAKSNKPHLDPPKRGPPQFGEVATATSKP SKKEYIGSRVAVEMALRGKSLMLIIGLGHGHPKKEFEISKXVHLDITDGRISTLETCT AGITPARIKTMEALMKKKP"	
gene	6254. 7306	

CDS	6254..7306	/gene="PF0316"
	6254..7306	/gene="PF0316"
		/note="Function Code: 16.1 Conserved Hypothetical"
		/codon_start=1
		/transl_table=11
		/product="hypothetical protein"
		/protein_id="A180440.1"
		/db_xref="GI:18892266"
		/translation="MKDIVESEVELLRKGMIEAKIVENIQENEDKVALCTIAEL TRDESELEAIYELRKIDHDKNAIGYEVAKYALDYEESLEFEBAVKIKMLN PVDKATLEIARLALAEVLDLEMPNAFEIIEIEMLEKIDLILOALEIA SGDALSENALKFEYMAODIYKMLKIGOSEGIIKEKLEIRLIMHTSADIKLAEV RFOALHEIEKPRGNNVPILEITPAWKIVPEELKIKETIMETSASKIDSEK KVVQLLEIDDLIKAYEIAITLESQELREKALVLTGKIMIDSEVAFVEKIFPLSD EKKEEFINFLKINPIA"
gene	7303..8145	/gene="PF0317"
	7303..8145	/gene="PF0317"
		/note="Function Code: 16.1 Conserved Hypothetical"
		/codon_start=1
		/transl_table=11
		/product="hypothetical protein"
		/protein_id="A180441.1"
		/db_xref="GI:18892267"
CDS		
Query Match	44.28; Score 825.2; DB 1; Length 13354;	
Best Local Similarity	66.68; Pred. No. 1e-151;	
Matches 1245; Conservative	0; Mismatches 608; Indels 15; Gaps 4;	
1	ATGACGGCGATCGAATGGAACACGAGACCTTTTCTAAGTCCCTACCTGGCGGACCGG	60
8179	ATGAGTTTCATCGAATGGAACGAAAAACGTTTGCAAGTTTCATCTGAGTGTATCC	8238
61	AGGATACGGGGAACCTTAATCGCGTACACCTTGACGAGGCGCAACATGAGACCAACAG	120
8239	AGGACAAAGGAGAAATGTGTGCTATGTTAAACAAAGCCCAACCTAAAGGACAAACAA	8298
121	TACGAGACGCGGTGTGTGTAAGACCTGAAAGCGCTCAAGGGCTTCATCGAGAAC	180
8299	TATGAAACCAATTTGTTATGGAACCTCAAGAACATCGAAGAGGTTTATGAAAT	8358
181	GCCTCAATGCCGAGATTTCCGACAGACGCGAAGCTCCCTTCACTGCTTTAAAGAG	240
8359	GCCACATGCGCAAGATTTCCCGAGTGAAGAAAGTACCTTTTGAAGGCTAATGAA	8418
241	GAGAGAGAGAGACCGAGATATGGTGGCCGATATCCAGACCTGACGCCCAAGAAATC	300
8419	GAAAAAGAGTAAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG	8478
301	CTCTCACTAAAGACGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	360
8479	CTTGAAAGCTAAGACATTAATGATCTTGGAAATGAGACGAACTCAAGAAATCTTCAT	8538
361	GTCGGCTTCAAGAGGAGGAGATGAGAGCTTCGCTTTGACGACAGCTCCCGCTGAG	420
8539	GTTGGCTTAAAGAGGAGGAGATGAGAACTTATATGAGAGATATGTTCAAGGCTG	8598
421	TTCGACAAATATGGGATCTTGTATGAGAGAGACGATCTTGCGCTTCTTGACACTGAG	480
8599	TTTGATGATCTTGATCTTGTATGAGAGAGACCAACATCTGAGATTTTGTGACAGGAG	8658
481	GCGGAGAGATATTCGAGATGCGAGAACGCGAGGTTTCAAGTGGCGCTTGCGAGCGG	540
8659	AGTGAAGAACTTATAGAGAGGTTGGAAGACCGCGCTTCTTCAGAGAAATTTGCAATGA	8718
541	GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	600
8719	GACAAATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	8775
601	TACGACATGATCTTATGAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	660
8776	TGGGACATCTATATCTGGGAGACGGAAGAAAGAAAGATGTTGAAAGATATCTGTT	8835
661	GAGGCGGTTGACTCCGACGGAAGAGAAATATCTCTGAGGCGGCAAGAAAAAGCGGTC	720
8836	TACGAGTATGATTCGATGAGAGAAAGATTTTGTCTATGAAAGCCGAGAAAGTAC	8895
721	ATCAGGACACGACGCTGATACCTCT-----GGAGCGGAGCTTAAACCAATCTAC	774
8896	ATGAGCGACCAATATAGCTTTATCATCTGATGAGAAAGATGAGGATCTCTGAT	8955
775	GAGGCGCCCTGACCTCTGGGAGACCAAGCTCACGGAAGAAAGGTCCTCTCACT	834
8956	GAAATGACAGAGAGATTTGGCAAGCAAGATATAAGATGGAAGATGATCTTACACTC	9015
835	CCAGATCGCGGACGAGGTTAACTCTGCTGCTGGAGCGGAGCGGAGCTTGTATAC	894
9016	TTTGAAGAGAGAGGCTTAACTATACATTTGGAGTGGGAAATTAACCCATAGCGAAG	9075
895	GCGGACCACTGATTTAGGGGCTTGACGTACGAGATGCGAAGCAATGCTCTCATCATG	954
9076	GGAAGCATTTGATTAATGAGATTTGAGT-----TGATGAGATGCTGTTTACCTTAAGAG	9132
955	ACCGCACGAGATAGCGAGCTCTACCTTACGAGACGCGACCTGAAACAGGTCAACGA	1014
9133	ACGGGCACTAAGATTAAGGAGGCTTACCTGGAGCGGAGAGAAACACTCACTGAC	9192
1015	TACACGGGCGCATTTACGAGAGCTCAAGACCTTGAGCCGAGCAGCTCCGCTTCAAG	1074
9193	TATTAAGATCAATCTTCTTAAAGATTTGAGAACTTTCAGGACGATCTCTGTTTAA	9252
1075	AGCAAGACCTCGAGATTAAGCGGCTGTACCTCACGCGGAGGATTAAGAG--GAGAG	1131
9253	AGCTGATGATTAAGATTAATGATGTTATGAAAGCCGAGCTCAAGGAGGAGAAAA	9312
1132	GCCCGGATATGATCTTCTGCAAGCGCGGCGGAGGAGGATGAGGACCGCTTGC	1191
9313	GCCCCGATATGATTTCTGCTGATGAGGCGGAGGAGGATGAGGATTAATCTTCAAG	9372
1192	TACGATGATGATGATGAGGAGGCTTACCTGCTGCTGAGACCGGCGGCGG	1251
9373	TATGAATGCAATTAATGCGGCGGAGGAGGATTAATGATTTAGCTTAACTCCAGGGG	9432
1252	AGCGAGGCTATAGCAAGACTTCCGCTCCGCTCTGAGAGAGAGCTGCTGAGAGAC	1311
9433	AGCAATGATATGATGAGACTTTCGACTAAGATTAACCAAGACTGAGCTGAGAGAC	9492
1312	TTTGAAGACATTAATGAGGAGCTGAGAGGTTCTTAAGCTGAGACCGGAGCGGACAG	1371
9493	TTTCAAGATATTAATTAACGATATGAGGATTCCTGAACCTTGAAGCGGAGGACAG	9552
1372	GAGCGGCTGATTAAGGAGGATTAAGCTTACGAGGCGGCTTATGACCAACTGGGCTTGA	1431
9553	GAGAGATGATGATGAGGAGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTA	9612
1432	CAGAGGACCTCTTCAAGGACGAGATTAAGGAGGAGGATTAAGGATTAAGGATTAAG	1491
9613	CAGAGGATGATGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTA	9672
1492	TACGCTTCTGAGACATTAAGGCTCTGCTGAGAGCTGAGAGCTGAGGAGGAGGAGG	1551
9673	TATGCAATTTTGGACATTTGCTTGTGATGATTAAGGAGGATTAAGGAGGAGGAGG	9732
1552	GAGAGGAGGATTAAGGAGGATTAAGGAGGATTAAGGAGGATTAAGGAGGATTAAGG	1611
9733	GAGAGGAGGATTAAGGAGGATTAAGGAGGATTAAGGAGGATTAAGGAGGATTAAGG	9792
1612	ATATCTTATATCACTGCTTGAAGACTACCGCTGCTGCTGAGGAGGCTTATGTT	1671
9793	ATTATATATATCACTGCTTGAAGACTACCGCTGCTGCTGAGGAGGCTTATGTT	9852
1672	TACAGGCTCTCAAGGATTAAGGAGGATTAAGGAGGATTAAGGAGGATTAAGGAGG	1731
9853	TATCAGGCTCTTGAAGGATTAAGGAGGAGGATTAAGGAGGATTAAGGAGGAGGAG	9912

	Matches	1086;	Conservative	0;	Mismatches	770;	Indels	32;	Gaps	7;
QY	1	ATGACGGCATCGAATGGAACCGAGACCTTTTCTAATGTCGCTTACCTGCGGACCCG	60							
Db	9370	ATGAGCAAAATTAATATGGGACGAGAGTAATCTTTACGAATTTGGCTTCTTAAAGACCCG	9311							
OY	61	AGGAT-----ACGGGGAACCTTAATGCGGCATACCGCTGACGAGAGGCCAATATAGAGC	114							
Db	9310	AGAAITTCAGAGGATGATTAATAATAGCTTATGTCTCTACAAAGGCGAGTTTGAAGAT	9251							
OY	115	AACAAGTACGAGACCGACGGTTGTTGTTGAAGACCTTGAAACGGGCTCAAGCGCTTCATC	174							
Db	9250	AATGAATACAAACATACCATTGTGTGAAA-----GAAAAGGAGACGAGATTTTAT	9197							
OY	175	GAGAACCCCTCAATGCCGAGATTTGCGACAGCGGCAAGAACGTCGCTTACCTGCTT	234							
Db	9196	GAAATATGCGAGATGCGCGCTTTCTCTTACGCGGAGGAAAAATGATTTTGTAAAGAG	9137							
OY	235	AACGAGAGAGAAGAGGACCGAGATATATGGTGGCCGATATCCAGACCTGAGCGCCAG	294							
Db	9136	CATAGGAGTAAGATGTTTAACGAGATATTTTATGCGATTTAAATTCATGAGTGAAGG	9077							
OY	295	AAAGTCCCTCACTAAAAAGCTCCGCTCGATGCAATGGAACGACGATTTCAAGAGACTC	354							
Db	9076	AAGATTTTGGAAATTAAGAGACATTTTGGATTTAACGTGGATTAAGATGACAGAGAAAT	9017							
OY	355	TTAATTTGCGCTTCAAGAGAGGAGACATGAGACCTCGATTTTGACGACGACGTCCG	414							
Db	9016	TTGATTAACAGGGAATTTAAAAAATGGAAGTATGAACCTCATATTTGAAGATGAGGTTCT	8957							
OY	415	GTCGTGTTTCACAATATGGAATCTCTTGATGAGAGAAAGACGACGCTTGCGTTCTGAC	474							
Db	8956	GTCGTGTTGATGATTAAGAGACATTTTGGATTTAACGTGGATTAAGATGACAGAGAAAT	8897							
OY	475	ACTGAGGCGCGAGAGATATATCGACAGATTCAGAGAACCGGAGTTTTCAGTGGCGCTCG	534							
Db	8896	GTGAGAGTGGAGATATCTTCTTGATGATGTTTCTCGAAGAAATTTTCTACGTCATCGG	8837							
OY	535	CACGGCGATCGATATGTTGTGACGTCCCGACCGCGAGGGGAGCAACCTTGCCCTGTC	594							
Db	8836	CATGAGATGAGTAATCTATTAATGTTCTTAAAAAGGTGGCAATGTTCCCAAGTTT	8777							
OY	595	AAGTTTACGACATGATGCTATGGAAGAGCGGGGAGAAAGAACTTTGAGAGGGCTC	654							
Db	8776	GGACTTTTGGCATATATACGTTTACGAGAGGAAAAAGAGAAACTGTTGAGGATGTT	8717							
OY	655	TCCTTCGAGCGGTTGACTCCGACGAGAAAGAAATCTCTGAGGGCGAAGAAAAAAG	714							
Db	8716	TCGATATACGTGTATGATTTTAACGAGAGATGGTGCTACTTTACGGTAACCCAGAAA	8657							
OY	715	CGGTTTCATCAGCGACGACGACTGGCTGTACTCTGTGGACG-----CGAGCTTAA	764							
Db	8656	GAAAAAGCACGACGAAATGATTTTCACTATCTGTGGGACGGGAGAAAGTTATCCCTTG	8597							
OY	765	ACCGATTCACAGGGCCCGCTCGACGTTGTGGAAAGCCAAAGCTCAGAGAAAGAGTCTA	824							
Db	8596	ACAGAAACTCTGTGTATACAAATACGACGAGGGGAAGTATGATGAAA--AGGAAATGTTA	8538							
OY	825	CTTCCCTCACTCAAGTACGGGCGAGGTAACCTCTGCGCTCTGGGACGGGAAAGCCGACG	884							
Db	8537	CTTTACTATGTCAAAAGAAAGGAGAAATTAAGCTTTTATAGCTGAGGAAAGAAATTTGGC	8478							
OY	885	TGTTGTTTACGGGCGACAC--TGATTTTACGGGCTTACGTCAGC---GATGGCAAGC	938							
Db	8477	TTCTATTTTCAGAGATTAATCATATGGGTAGGAAATTTTACAGCTTACGCGAGAGGAAAGT	8418							
OY	939	ATTGCTCTCATCATGACCGCCACAGAGATAGGCGAGCTTACCTTACAGAGCGGAGCT	998							
Db	8417	GGTTTTCACATAAGAGACTGATATTTATCCAAAGCAACTCATGTATGCGATGAAGGCT	8358							
OY	999	GAACAGGTCACCGAATTAACAAGGGGCGGATTTTCAGAGAGTCGAAGACCTTCAGGCGAG	1058							
Db	8357	TGACACAGTGCAGATTAACAATAAGCTATGATGATGATGAGCTTGATGTGAAGCATATTA	8298							

QY 1059 GCATCTCCGCTTCAAGACCAAGACCTCGAGATAGACGGCTGTACTCAGCCGAGGT 1118
 DB 8297 CAATTTAAATACAAAGCTGATTGATTAATGATGATGATGATTAACCTGAAAT 8238
 QY 1119 TAAAGAGAG--AGGCCCCGGTATAGTCTCTCCACGGCCGCGGAGGATGTA 1175
 DB 8237 AGAAGAGGCGAAAGGCTCCGGTTTGGTTTGTTCATGCGGCTCCAAAGGGGATGTA 8178
 QY 1176 CGGACACCGCTTCTGCTACGAGATGACAGTGTGCGGAGCAAGGCTCTACTGCTGCT 1235
 DB 8177 CGGGATATTCTTTAAATACGAATTTCAACTTATGCGAATAGAGATTTTATGTGTTAT 8118
 QY 1236 CGTGAACCCGCGCGGAGCGAGCGGCTATAGCGAAGACTTCGCGCTCCGCTCTGAGAG 1295
 DB 8117 TGTAAATCTAGAGGAGATTAATGCTACGCGCAAGATTTCCATGTAGAGTGTGGAAG 8058
 QY 1296 GACTGCTTGGAGACTTGTGAGCATATGAAAGCATGAGAGTCTTCAAGCTCGA 1355
 DB 8057 AACAGGCTAGAAAGTTTTCAGACATATTTAAATGGGTAGAAAATTTCTTGAATTA 7998
 QY 1356 ACCGAGGCGCAGAGGAGCGCGTGAATTAACGGGCTAAGCTACGGCGCTCATGAC 1415
 DB 7997 ACCAAGCGCTATAGAGAAAGATTGGAATTAACGGGTATAGCTATGTTGGTTATGAC 7938
 QY 1416 CACTGGGCTTGTACTCAGACGACTTCAAGGCAAGATTAAGCCAGAAAGCATTAAG 1475
 DB 7937 TAAATGGCTTGTACTCAGACGACTTCAAGGCAAGATTAAGCTGCAATTAAGTGAAT 7878
 QY 1476 CTACTGGCTACACGCTACGCGCTTCTGAGCATAGGGCTGTGATGACGCTCGAGTAT 1535
 DB 7877 TTACTGGCTTACAGATTAAGCTTTTCTGACATGATTAATGTTTACAGAGATGAT 7818
 QY 1536 CGGCGCAATCCGTTAGAAAGCAAGAACTTCAAGAGCTGACCGGCTGTCTACGCTCA 1595
 DB 7817 TGGGGAATATCCACTTGAATTAATTAATTAAGAAATTAAGCAATTTTTCAGCCAA 7758
 QY 1596 GAAGGTGAGGCGCGGATGCTACCTCAATCCACTGCTGAGACATACGCTCTCGCTCGA 1655
 DB 7757 AAATGTGAGAGCCCACTCTTGTGATTCATAGTTTGAAGACATCAAGATGCTCCCTGA 7698
 QY 1656 CCAGAGGCTTATGTTCTACACGCTCTCAAGAGCATGGGCAAGGCTTACATAGCAT 1715
 DB 7697 CCAGAGGCTATGTTTATACGCTGCTCAAAAGCCTTGGCAAAAGCCTTATAGCAT 7638
 QY 1716 ATTAAAGCGGCGCGCCACAGCGCTCCGCGAAGCCGAGCAACAGCCGGAAGCG 1775
 DB 7637 CTTCAAAAAAGAGAGCTCAGCGGCGACAGCTGTGATGAATCTGCTCAGAGGCAAAAAG 7578
 QY 1776 CTACAGGCTCTTCATAGTCTTCGAGCGCAAGCTCAAGAAAGTACAGAGGCTTTGA 1835
 DB 7577 GTATAAATCTATTTTGAATTTTGAAGAAAGTCAAGATATGAAGAGGTTTGA 7518
 QY 1836 GGTAGAGAGATTAAGGGAATGGG 1863
 DB 7517 TGTGAGGAGGATTTTGAAGAGGAG 7490

RESULT 9
 LOCUS AP000064/c 247695 bp DNA linear BCT 06-APR-2000
 DEFINITION Aeropyrum pernix genomic DNA, section 7/7.
 ACCESSION AP000064 BA000002
 VERSION AP000064.1 GI:5105945
 KEYWORDS
 SOURCE Aeropyrum pernix (strain:K1) DNA.
 ORGANISM Aeropyrum pernix
 Archaea: Crenarchaeota: Thermoprotei: Desulfurococcales;
 Desulfurococaceae; Aeropyrum.
 REFERENCE
 AUTHORS 1 (sites)
 Kwarabazasi, Y., Hino, Y., Horikawa, H., Yamazaki, S., Halkawa, Y.,
 Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankel, A., Kosugi, H.,
 Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Nakazawa, H.,

COMMENT

TITLE JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Tanaka, T., Hino, Y., Kwarabazasi, Y. and Kikuchi, H.
 Submitted (14-DEC-1998) Yutaka Kwarabazasi, National Institute of
 Technology and Evaluation, Biotechnology Center, Nishihara 2-49-10,
 Shibuya-ku, Tokyo 151-0066, Japan (E-mail: kyutak@kazusa.or.jp,
 Tel: +81-3-3481-8951, Fax: +81-3-3481-8424)
 Kwarabazasi, Y. is officially affiliated with the National
 Institute of Bioscience and Human Technology, Tsukuba, Ibaraki
 305-0046, Japan.
 Nomura, N. and Sako, Y. are at Kyoto University, Sakyo-ku, Kyoto
 606-8502, Japan
 The other authors are at the National Institute of Technology and
 Evaluation, 2-49-10 Nishihara, Shibuya, Tokyo 151-0066, Japan.
 All the sequence with length 100 codons or more between ATG or GTG
 and stop codon are defined as CDS
 Homology analysis is performed by Smith-Waterman algorithm against
 Genbank and GenPept release 109; EMBL release 56.0; SwissProt
 release 36.0; PIR-Protein release 57.0; and OMT release 31.0.
 E-mail address for comments and questions: genome@nitech.go.jp
 Restriction map, ORF organization, sequence alignment and more
 information are available at W.M.W. site of Biotechnology Center,
 URL: <http://www.mild.nitech.go.jp/>.
 Location/Qualifiers
 1..247695

FEATURES

Source

/organism="Aeropyrum pernix"
 /strain="K1"
 /db_xref="taxon:56636"
 complement(721..798)
 /gene="APECRNA43"
 complement(721..798)
 /gene="APECRNA43"
 /product="rRNA-His"
 /note="codon recognized: CAC; anticodon: GTC"
 /anticodon="(pos:761..763,aa:His)
 complement(980..1501)
 /gene="APE2245"
 complement(980..1501)
 /gene="APE2245"
 /product="Ape2245"
 /codon_start=1
 /transl_table=1
 /product="173aa long hypothetical protein"
 /protein_id="BA81257.1"
 /db_xref="GI:5105946"
 /translation="MLNDVSIIEVNDPEAGVPCVCHNIPMASPROPLATHNIRPSP
 SLGEVGLDEPPRRRAGGEGVCEVLEFREGDGRVRVNPVPPRPIAAGQWQD
 VSNAAPEPEALOSTIPGPGVTPAPRRPPHNNRPHRDEIPRLAQPPRTVAQIDL
 PNOPTPTPTIHNH"
 983..1549
 /gene="Ape2246"
 983..1549
 /note="motif-eukaryotic putative RNA-binding region RNP-1
 signature"
 /codon_start=1
 /transl_table=1
 /product="188aa long hypothetical protein"
 /protein_id="BA81258.1"
 /db_xref="GI:5105947"
 /translation="MYVDGKGCWLVSEVYTTIGSGSLGELIDLVESGSRVVDVR
 RMPRSRRPWASRDLFGRGKVAITLIPALGGRSFRGVDVPSIRALFSCFESG
 FNAIYAIASPPAMRGLIYAIYLAEGAPVWCSERLIPWRKRVADMLYARGFR
 VLHLDSDSIYERGTGCAAMLGSPSSV"
 complement(1536..2120)
 /gene="Ape2247"
 complement(1536..2120)

CDS

gene

	gene	/gene="APE2252"	/note="similar to OML:AB00947119 percent identity:51.042 in 288aa."
	CDS	/codon_start=1 /transl_table=11 /product="303aa long hypothetical maltose transport system permease protein" /protein_id="BA81264.1" /db_xref="GI:5105953" /translation="MMSGNPLEEVVSGSISRRIKIVIALALLAGSPILLTYGLILS SFSTNWDPLDLPQPTLDNNIMFLRGGITATAGGLENNVEMLVLTVAVGALT AETLVSLAAVSISRMKEFGKGVFTAGLIMHAFPGVALIGVMIFRIGIDVLGYGE EFAFVAIYIIARALAEVPVAILMEFKGFPKIPMELEMSSALVGAASRVRRMSIIVPLV KPIIAVLITFFLAGMEEFIFYFYFLPKPDISLPFFIEIYVAAMENQLTIIAAGT FYLLPIYTPVFYTORLLERYSGLKG"	
	gene	complement(5237). .6067) /gene="APE2253"	
	CDS	complement(5237). .6067) /gene="APE2253" /note="similar to OML:AP000001208 percent identity:59.058 in 276aa." /codon_start=1 /transl_table=11 /product="276aa long hypothetical lactose transport system permease protein" /protein_id="BA81265.1" /db_xref="GI:5105954" /translation="MRNMNVDAITFDIYGLMYKYLYHPAKYDPERNALITTVFA VTIVVVNLGGITRLAITFPFEIERISIHYYLMLLRMPVANYSLNTYFFPSDLGT LNKLILAGYIDKFINMGSTMPLGWMTLITVNGLVGSFSEIYFSIKRSIPSETI AARVVGASNLVIIRRIIVPMIRMHLFVTWQMSSLTTYTHIFLLEWGVDFTWG QTMSLIYFNLAFLPYVELROGLAAAAATLVYIGAIIIGLVTLRIIGFORMITRPGRDI "	
	gene	complement(6197). .7948) /gene="APE2254"	
	CDS	complement(6197). .7948) /gene="APE2254" /note="similar to OML:AB009472 percent identity:52.874 in 522aa." /codon_start=1 /transl_table=11	
Query Match	Best Local Similarity	11.08; Score 204.8; DB:1; Length 247695;	
Matches 541:	Conservative	0; Mismatches 452; Indels 15; Gaps 5;	
Dc	811 GAAGAAGAGCTACTCTCCTCACTCCAGATGCGGCACAGGATAAACCTCTGCGTCGCGAC	870	
Dc	42085 GGAGGAGAGCGCTGGTAGTGCTATGACAGACAGGCGAGAGCGTTATACGCCCGCGGMA	42026	
Oy	871 GGAAGAGCCAGCGCTGTGTTACCGCGGACCACTGCATTACGGGCTTGACCTCACGAT	930	
Dc	42025 GGACGGGTTAGAGCACTAGTTAAGGGGGGATATGATGTTGTCAGGCAACATCCATGGA	41966	
Oy	931 GGCAAACATTTGCTCCATCATGACAGCGGACAGAGATGAGGAGCTTAC---CTCTAC	987	
Dc	41965 GAGAGAGTTGTGTRCAATAGCTCTCTCAACAAGGACCCCCGAGAGTGTACAGGTTTAAG	41906	
Oy	988 GAGCGGAGCTGAACAACAGCTCACCGCAATACAGCGGCGCATATTCAGAAAGCTCAAGACC	1047	
Dc	41905 GCGCGGAGTGTGAGAGGGGTGTCAACGATTAACCGCTGGTGTCTCAGAGAGTTCAGACTG	41866	
Oy	1048 TTGAGAGCCGAGGCACTCCGCTTCAAGAGCAAGAACAGCTCGAGATPAGAGCGTGTACCTC	1107	
Dc	41845 TACAGCCCGCGAGAGGTCTCAAGTAGAGGCGGAGGCGAGGTTGTCCAAGGGCTGTGTRACATA	41786	
Oy	1108 AGGCCGAGGTTAAAGAGAGAGAGGCCCGCGGTATGATCTTCGTCACAGCGCGGCGGAG	1167	
Dc	41785 GAGCCGAGGTTGAGGGGAGAGAGGCC---TAATACATTAATACACGAGAGTCTCTAAG	41729	
Oy	1168 GGCATGTACGACACCGCTCTGCTCATGAGATTCACACTGATGGCGAGCAAGGGCTCTAC	1227	
Dc	41728 GGCATGTACGACCTATGAGATTCACGGCGAGATTCACCTCTTCGCATCCAGAGGCTTCACA	41669	

OY	1228	TGCTGCTGCTGACACCGCGCGGACGACGCGCTATAGCAGACCTTCGCGCTCCGCTC	1287	TITLE	Direct Submission
Db	41668	GTGGCTCTACGCGACCGGAGGCTGACGAGGGTATAGAGAGAGTTCGCG---CGACATA	41612	JOURNAL	Submitted (08-NOV-1999) The Institute for Genomic Research, 9712
OY	1288	CTGAGAGGAGCTGCTTGGAGGACTTTGGAGACATATATGAGCGGATCGAGGAGTCTTC	1347	FEATURES	Center Dr, Rockville, MD 20850, USA
Db	41611	AGAGGAGGATATGGGATATACCGACTACAAACAGCTATAGATTCCTGGACGCTGTTGTG	41552	source	1. 10256
OY	1348	AAGCTTGAAACCGAGCGGACGAGAGGAGCGGTTGGAATAAGCGCATTAAGCTACGGCGG	1407	/organism="Deinococcus radiodurans"	
Db	41551	GAG---GATAGCATCTGTGACGAGGATAGATGGCGGTGACATAGCTACGGCGG	41495	/strain="R1"	
OY	1408	TTCAATGACCACTGGGCGCTTGTGACTGACGAGACCTTTCAAGGAGGAATAGCGGAAC	1467	/db_xref="taxon:1299"	
Db	41494	TACATGACAAAGCTCATCTGATACGAGAGAGAGAGGATTCAGACACTGTGAGCGGAAT	41435	/chromosome="1"	
OY	1468	GGCATTAAGCTACTGCTGCTACACGACGCGCTTCTCGGACATATAGGCTGTGTCAGCGTC	1527	/gene="DR0164"	
Db	41434	GGCATTAAGCTACTGCTGCTACACGACGCGCTTCTCGGACATATAGGCTGTGTCAGCGTC	41375	/note="similar to GB:D14846 P1D:286165 percent identity: 49.11; identified by sequence similarity; putative"	
OY	1528	GAGGTGATCGGGGCGCAATCCGTTAGAGAA---CGAGAACTTCAGAAAGCTCAAGCGCTG	1584	/codon_start=1	
Db	41374	GACCAATATAGCGGCGACCTCCCTCGACAAATCTGAGAGATAGCTTAAAGAGAGCGCGGA	41315	/transl_table=11	
OY	1585	TCTCAGCTCAGAAAGCTGAAAGCGCGGATACCTCTAATCAGCTCGCTTGAGGACTACCGC	1644	/product="endo alpha-1,4 polygalactosaminidase, putative"	
Db	41314	TTTCACGCAAGAAAGCTGAAATCTCCAGTAATGCTAATTCACAGTATGAGAGATTCACCGC	41255	/protein_id="AAF09753.1"	
OY	1645	TCTCGGCTCAGCAAGAGAGCTTATGTTCTCTCAACGCTCAGGAGACATGGGCAAGAGAGC	1704	/db_xref="GI:6457833"	
Db	41254	TGCTTCATATAGCAAGCGCGCTCGCAATGACACAGCTCTTAAGCTCAACGCGCAAGAGAGC	41195	/translation="MRIPLTVLPLSAVLLASGCTPSPVAATVPATPAOPTRTLA	
OY	1705	TACATAGCATATTCAGAGCGCGCGCCACAGCGGACAGCGTCCGCGGAAGCCGAGAGC	1764	AAATPLRLPPLPAGKLAADMOIGAAATGAKYALPGVSLDLDGFTSAAVADLKAGV	
Db	41194	TTCTAGTGTGTCTTCACCAAGGAGAACCGCGGACAGATATCTCGCTCGCGGCTCAC	41135	YVYCYLVNGSYESYRPAADAOVPSIKITQIDPMPDSEFYDIDVREGGVLLDGR	
OY	1765	AAGCGCAAGCGCTACAGGCTCTTCATAGAGTTCTTCGAGCGGACGCTC	1812	IALCAKGFDAVPRDLNDQNTVTSVIRQDLDNGSLADRAHAGLAILAKNPD	
Db	41134	AGGAGGAAGAGCTCGAGCTCAAGCTCAAGTGAATAAAGAAAGAGCTC	41087	YVLAADRGRLVLDVTVNESQRTKCGPLTEYVRKGLALINVEYKADLNCAITH	
RESULT 10				NSLGVNALLFRDLSLVKRLQKAYKRVSCG	
AE001879/c				/complement(1079. 3046)	
LOCUS	AE001879	10256 bp	DNA	1linear	BCT 22-NOV-1999
DEFINITION	Deinococcus radiodurans R1 section 16 of 229 of the complete			/gene="DR0165"	
ACCESSION	AE001879	AE000513		/note="similar to GB:A1009126 percent identity: 56.01; identified by sequence similarity; putative"	
VERSION	AE001879.1	GI:6457832		/codon_start=1	
KEYWORDS				/transl_table=11	
SOURCE				/product="acyl-peptide hydrolase, putative"	
ORGANISM	Deinococcus radiodurans.			/protein_id="AAF09754.1"	
	Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;			/db_xref="GI:6457834"	
	Deinococcaceae; Deinococcus.			/translation="MNNSETPAPGPDLSLALAPSPDPSPGKQVAFVLAIDSEDP	
REFERENCE	1 (bases 1 to 10256)			AKFDKDFARTRISGLMLSEGGARPLTAETGKGSAPRMSPDQONLAFVSADEVK	
AUTHORS	White, O., Eisen, J.A., Heidelberg, J.F., Hickey, E.K., Peterson, J.D.,			AAALMLPLLGGEARVTHFRKNGVSGPQSPDGRFLFTTATLWTEPRKRDREARVLT	
	Dodson, R.J., Haft, D.H., Gwinn, M.L., Nelson, W.C., Richardson, D.L.,			PVYRANGADMLPERPALMLYDEADKLEWYAEIIGALSMPPDSRGLVLIQSEDE	
	Moffat, K.S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M.,			WQASORQDVYDLPPLPADAPAPKQLMNSAAGLAPHPGQRFALGRAGNGT	
	Vamathavan, J.J., Lam, P., McDonald, L., Uitterback, T., Zalewski, C.,			EAHATLTIENGQRRRLDGHDPADVAGGDCVGAAPRGPRLDGDPTLPSSTVRS	
	Makarov, K.S., Aravind, L., Daly, M.J., Fraser, C.M. et al.			VGCFPAHIGGCKRAVDHPQGVSAFTANDEGVALRESATRPPEYELVNGQVYTDLA	
	Genome sequence of the radioresistant bacterium Deinococcus			RPFPVREPDQVTFETELDEGEGWVLPPEGEKVAPLLIHGGPRTDVGHGFTHEFL	
	radiodurans R1			AAVGSYSGEMTWITGTRFOAATIDRCLSNLSFGSTSDIGLRFMDDELGDFSR	
TITLE	Science 286 (5444), 1571-1577 (1999)			RADALKMLDSPLOYEVNKTPLIYHVSILDRHCPPEQAEQVYALHKKHQVYVRYFR	
JOURNAL	20036896			PERNHLSRSGRDRRLTRLNETFANLEML"	
PUBMED	10367266			3098. 3352	
REFERENCE	2 (bases 1 to 10256)			/gene="DR0166"	
AUTHORS	White, O., Eisen, J.A., Heidelberg, J.F., Hickey, E.K., Peterson, J.D.,			3098. 3352	
	Dodson, R.J., Haft, D.H., Gwinn, M.L., Nelson, W.C., Richardson, D.L.,			/note="similar to SP:P31787 P1D:1323021 P1D:1620454	
	Moffat, K.S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M.,			P1D:170970 percent identity: 66.09; identified by sequence	
	Vamathavan, J.J., Lam, P., McDonald, L., Uitterback, T., Zalewski, C.,			similarity; putative"	
	Makarov, K.S., Aravind, L., Daly, M.J., Minton, K.W.,			/codon_start=1	
	Fleischmann, R.D., Ketchum, K.A., Nelson, K.E., Salzberg, S.,			/transl_table=11	
	Smith, H.O., Venter, J.C. and Fraser, C.M.			/product="acyl-CoA-binding protein"	
				/protein_id="AAF09755.1"	
				/db_xref="GI:6457835"	
				/translation="MFEQAGQEVNTLSKPPGNDVILKILXLYKQGSAGDVGKRRPG	
				FDVVGAKYDAMWGLKGTQEOAQDEYVALVEELAKKQDQ"	
				3663. 4649	
				/gene="DR0167"	
				3663. 4649	
				/note="identified by Glimmer2; putative"	
				/codon_start=1	
				/transl_table=11	
				/product="hypothetical protein"	
				/protein_id="AAF09762.1"	
				/db_xref="GI:6457842"	

RESULT 11
CNSPAX06/c

LOCUS 265118 bp DNA linear BCT 09-SEP-1999

DEFINITION Pyrococcus abyssi complete genome; segment 6/6.

ACCESSION AJ248288 AL096836

VERSION AJ248288.1 GI:5458960

KEYWORDS complete genome.

SOURCE Pyrococcus abyssi.

ORGANISM Pyrococcus abyssi.
Archaea; Euryarchaeota; Thermococci; Thermococcales;
Thermococcaceae; Pyrococcus.
Hollig, R.
Pyrococcus abyssi genome sequence: insights into archaeal
chromosome structure and evolution
Unpublished
2 (bases 1 to 265118)
Genoscope.
Direct Submission
Submitted (06-JULY-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref.genoscope.cns.fr
- Web : www.genoscope.cns.fr)

JOURNAL TITLE
AUTHORS
REFERENCE

FEATURES
source

gene
1
CDS

CDS

CDS

CDS

CDS

/organism="Pyrococcus abyssi"
/strain="Orsay"
/db_xref="taxon:29292"
164..1615
/gene="bgal"
164..1615
/gene="bgal"
/note="PAB2376"
/codon_start=-1
/transl_table=11
/product="BETA-GALACTOSIDASE (EC 3.2.1.23) (LACTASE)."
/protein_id="CA50447.1"
/db_xref="GI:5458961"
/translation="MEFTWGVVQSAQFENGDDPYRRNLDRSDMHWVRDYPNKNEL
VSUDELDEGINNELLEYEDIRRLAKELGLNLYOULIEMSRTEPCPTVEVEEVERDGCG
FISVARIKREHELEKLDELANKRERHYENYLKMLKINFTEVTALNTGPIMLHDPI
EVVAIVLEKAAARGWVERAIAVEFSKAAYAAKMDKYDFMAFEDEPMVAEGUYAP
YVMPGILNPRAKASVIMHOITAIHARAVIDSTIKESDPKPGVILNITIPADLPBRSK
HYAAANYDLGFHNRLFLFAANKRGKDIDINGEYKKYPHLKRNDTGNNTTYREVYKV
EPYKELPLTVFYGVBEIGYTSANPNVISPDNNPTSDFGWEVYPRGLDYDAEAVERSE
MWETENGDIADSKDILRPRIYVDHVKEVKRLRNGLNVLGYFHVALTDNYENMAGFERI
REGELYEDPTIKERIIPRKRSVETYYKKVKGEGIE"
1834..2283
/note="PAB1019"
/codon_start=-1
/transl_table=11
/product="hypothetical protein"
/protein_id="CA50448.1"
/db_xref="GI:5458962"
/translation="MPLPIDTFDSVTLKMKHSEKNRRLLETLRKFTYVLSMLTYOR
VTAIEGRYLRNVLEFEFVGLKIDINIYVSEILLTRASKTEANLRGVFDLEDIIITA
YLAIEGSLIIIDPEPKYEPIRRFRGLDTNPDLKFLKREVELAKEKITI"
complement(2276..3079)
/note="PAB1341"
/codon_start=-1
/transl_table=11
/product="hypothetical protein"
/protein_id="CA50449.1"
/db_xref="GI:5458963"
/translation="MFENLESFEAFKLLIEVLERDLNDPFQDVKEVICIEHPDDC
VIWGMYRTIKLEERGIELEVYAACMTDGMGTLDSSLGHTELATIRREEEBSSTLVK
KIYMLVGRDTLEPYSEREVARDLVAIRIKREPQDVFPLPYEARPHDNBTGFLAD
AAAFSPNPNVNDVIEIGPHOVSFIALITYNNKPVPNFVDIIVYMELKKAITHHSO
PPDDVWEVNEPLFRVAFYGGKAGVYAGDFRIMPGLTFYHIITPADLI"
3192..4340
/note="PAB1020"
/codon_start=-1

CD5

/product="hypothetical protein"
/protein_id="CAB50450.1"
/db_xref="gi:5458964".
/translation="MKEVSVYKELIVKLGITREDRIETLEMKGIIIEFDGIRYR
FDSACKLRGTVGIDEEVYIPGFPHIRIINLSRIGIRIKRGEFYEEKVGQYR
VYMKKMGITRGVIGIECFETTERIPDPVPOEFKDNPLIVGEMGSPSYVEEP
PYKEDIOGFLEFDVOIKGRSLPYEELKIAEEGIHNVPECKYIKDDVDELXOI
I ERLKEDGFLIDKSPDMKIKYKVTYTPANINDIKGARYVELPRGYFSRIRLAF
YLAERIKKEEERERAKELGSLDPPVESIFDVQEDHLELRVKKYKRETRAKV
THEKLGLEIIVDIEIINDGWRITFKRLYPDAINEIRELIGKAFV"
4351..5658
/note="PAB1021"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB50451.1"
/db_xref="gi:5458965".
/translation="MEDIALRLVKAIRADPYKAVINSVRENGKLLVKKDEFEIG
KYVLAFGACDMARAIIEILIDVEDGAVYKGYGKELKIRVIEAGHPIDPNSIR
GALIEGKILEKYEEDIVFYLISGGSGALFKPEEGISLEBLKNTLILKSAKTHE
INMYRKHISYKGGKGLAKIKKGQYVILISVDNENLAIASGPTVDPPTPODKARI
LELYGMHVPSPVSRSYHTEKIRGEAEETLEBELPNHNLILANSISCAVYBEAK
LDFKNAIMVTVLLEGAKDAGLEIGSIVOEIAKGNPPEPVVLVFGGETVTVLIEGGG
KGGPMDLAIATSTRKIKGLNAILVAFDPDGDGPTDAGGVLDGETYEKLRGIDIE
KYLKENSYEALKKYGSLFTFGPTGNVNSMIILITSKOGR"
complement(5633..6655)
/note="PAB1338"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB50452.1"
/db_xref="gi:5458966".
/translation="MRIYEARKVGISSSWVTSKYSNRDALAOLSDRIYVKTLRNGE
VYGESEPSALTDIGIRIPDEIKINPEKDHGKFLVPRGELVETIEEMLIL
EKKEPEVYKHYEVLINRKYTMLOLARYGCAINMESWEGHLYIKVNASATQTE
RSVYVITDOKRVSITFSDLEDEIEEYEMSKRKYRANKIRFPHNSVTSILIPKA
VOLVETVILDKRTPAMEFIIKIADEFPYKSEOELEMELELSDMEQOIIIV
LYSGNIPLEHQFLDITEKEIEEYDRMIDKLKLYIMIRIVDTLNEGKRIYVKKLY
YGLVSN"
complement(6670..7425)
/note="PAB1337"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB50453.1"
/db_xref="gi:5458967".
/translation="MDVGTILILAVVYLIGSIVASIVSSVISOKMNMALINSL
EYETVYKKRKGKRGSGSARKEPVNAKQVRSVEDYVYRKIOTIDEXHANNANESK
KKERLIMLDVARGTALGFSEDEINAFKMLVSELDERFLMLAIPPEPDIDNRKK
MLTFAKRIPIEYERKKGKGEISVKSPPDALLRITSSIKSAISILDILIKRGNP
ATVPEIKLSOLEKCKKRYEDLEKRLUDELEAL"
complement(7468..9687)
/note="PAB136"
/codon_start=1
/transl_table=11
/product="methyl-accepting chemotaxis protein (t)pc"
/protein_id="CAB50454.1"
/db_xref="gi:5458968".
/translation="MQFRKILILIASAVITMAIFGVTISFNKRYGVYKGEIOAP
AIEGGGLAMSAEAYEAKEMDFDFAVADGATATVEYKAYAGKIGDSNTEAREIL
SEFKRIKINDVSVYVYGGDEGHWYAPDELPEDGDPYRVRVYQAKEGALITTE
PYRDAATGQVITVAPVYVNGQKRYGIDLVPSYTLNFKKRIKLRKTYIATTON
GATLVHPDEEITINKLIPQVELKDLDLCKELGDESEVLYTYQGIKIGGKRIKLTQ
GMIVAVVAPLSDIGPILISADTDAIINKMKIMVIOGFAVGVIIAIVLIVYKASWT
APLEKTAQAALDAEGRILQVSEYKIOIRLEIREGALIOAEVASKDVGJNANIS
KLESLADGSLNIGVEYREGELRDIIEDIRSYTFRSGISGLVMADLEKRAAL
AOVSDYVEALINQVNAIIOQVSIQABROQETINDEIGMRLVAKOTSSESVLMEERG
AVTEVYIANGSOKDBALKRIEDIOHMSRIEETYSKAEKSNRIEETIVYTSIA
EOTILNANLAIENARGEARGFVAVAOETRIKLAESKOADNIKSIIIDKTLDEIK
AVEATKRESVYISGESSELTQDTIGYLANIATLLOETSERMTYKQIVPRQEVKAL
RALENILASAEETTASAEVSSAIEQOTAAIEELRRAADEILKMDVGRMROIGKRYV
complement(9704..10195)

CD5

/product="hypothetical protein"
/protein_id="CAB50455.1"
/db_xref="gi:5458964".
/translation="MKEVSVYKELIVKLGITREDRIETLEMKGIIIEFDGIRYR
FDSACKLRGTVGIDEEVYIPGFPHIRIINLSRIGIRIKRGEFYEEKVGQYR
VYMKKMGITRGVIGIECFETTERIPDPVPOEFKDNPLIVGEMGSPSYVEEP
PYKEDIOGFLEFDVOIKGRSLPYEELKIAEEGIHNVPECKYIKDDVDELXOI
I ERLKEDGFLIDKSPDMKIKYKVTYTPANINDIKGARYVELPRGYFSRIRLAF
YLAERIKKEEERERAKELGSLDPPVESIFDVQEDHLELRVKKYKRETRAKV
THEKLGLEIIVDIEIINDGWRITFKRLYPDAINEIRELIGKAFV"
4351..5658
/note="PAB1021"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB50456.1"
/db_xref="gi:5458965".
/translation="MEDIALRLVKAIRADPYKAVINSVRENGKLLVKKDEFEIG
KYVLAFGACDMARAIIEILIDVEDGAVYKGYGKELKIRVIEAGHPIDPNSIR
GALIEGKILEKYEEDIVFYLISGGSGALFKPEEGISLEBLKNTLILKSAKTHE
INMYRKHISYKGGKGLAKIKKGQYVILISVDNENLAIASGPTVDPPTPODKARI
LELYGMHVPSPVSRSYHTEKIRGEAEETLEBELPNHNLILANSISCAVYBEAK
LDFKNAIMVTVLLEGAKDAGLEIGSIVOEIAKGNPPEPVVLVFGGETVTVLIEGGG
KGGPMDLAIATSTRKIKGLNAILVAFDPDGDGPTDAGGVLDGETYEKLRGIDIE
KYLKENSYEALKKYGSLFTFGPTGNVNSMIILITSKOGR"
complement(5633..6655)
/note="PAB1338"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB50457.1"
/db_xref="gi:5458966".
/translation="MRIYEARKVGISSSWVTSKYSNRDALAOLSDRIYVKTLRNGE
VYGESEPSALTDIGIRIPDEIKINPEKDHGKFLVPRGELVETIEEMLIL
EKKEPEVYKHYEVLINRKYTMLOLARYGCAINMESWEGHLYIKVNASATQTE
RSVYVITDOKRVSITFSDLEDEIEEYEMSKRKYRANKIRFPHNSVTSILIPKA
VOLVETVILDKRTPAMEFIIKIADEFPYKSEOELEMELELSDMEQOIIIV
LYSGNIPLEHQFLDITEKEIEEYDRMIDKLKLYIMIRIVDTLNEGKRIYVKKLY
YGLVSN"
complement(6670..7425)
/note="PAB1337"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB50458.1"
/db_xref="gi:5458967".
/translation="MDVGTILILAVVYLIGSIVASIVSSVISOKMNMALINSL
EYETVYKKRKGKRGSGSARKEPVNAKQVRSVEDYVYRKIOTIDEXHANNANESK
KKERLIMLDVARGTALGFSEDEINAFKMLVSELDERFLMLAIPPEPDIDNRKK
MLTFAKRIPIEYERKKGKGEISVKSPPDALLRITSSIKSAISILDILIKRGNP
ATVPEIKLSOLEKCKKRYEDLEKRLUDELEAL"
complement(7468..9687)
/note="PAB136"
/codon_start=1
/transl_table=11
/product="methyl-accepting chemotaxis protein (t)pc"
/protein_id="CAB50459.1"
/db_xref="gi:5458968".
/translation="MQFRKILILIASAVITMAIFGVTISFNKRYGVYKGEIOAP
AIEGGGLAMSAEAYEAKEMDFDFAVADGATATVEYKAYAGKIGDSNTEAREIL
SEFKRIKINDVSVYVYGGDEGHWYAPDELPEDGDPYRVRVYQAKEGALITTE
PYRDAATGQVITVAPVYVNGQKRYGIDLVPSYTLNFKKRIKLRKTYIATTON
GATLVHPDEEITINKLIPQVELKDLDLCKELGDESEVLYTYQGIKIGGKRIKLTQ
GMIVAVVAPLSDIGPILISADTDAIINKMKIMVIOGFAVGVIIAIVLIVYKASWT
APLEKTAQAALDAEGRILQVSEYKIOIRLEIREGALIOAEVASKDVGJNANIS
KLESLADGSLNIGVEYREGELRDIIEDIRSYTFRSGISGLVMADLEKRAAL
AOVSDYVEALINQVNAIIOQVSIQABROQETINDEIGMRLVAKOTSSESVLMEERG
AVTEVYIANGSOKDBALKRIEDIOHMSRIEETYSKAEKSNRIEETIVYTSIA
EOTILNANLAIENARGEARGFVAVAOETRIKLAESKOADNIKSIIIDKTLDEIK
AVEATKRESVYISGESSELTQDTIGYLANIATLLOETSERMTYKQIVPRQEVKAL
RALENILASAEETTASAEVSSAIEQOTAAIEELRRAADEILKMDVGRMROIGKRYV
complement(9704..10195)

```

/note="PAB1335"
/codon_start=1
/transl_table=11
/product="CHEMOTAXIS PROTEIN CHED"
/protein_id="CAB50455.1"
/db_xref="GI:5458969"
/transl_table="MREIKVIGIDYANGKGGIISTVYSGSCVGLTYDRYKXGCL
LHALPEARICRGKPNKPYDTGDLLEKYLKGLASFRLEAKLFGAGQFONIKS
DELKIGERNVOTAKRRLKRLGIRLVAEDTGGGRTIYDLSGTGRMKRVIGGOVE
KY"
complement(10179..10817)
/note="PAB1334"
/codon_start=1
/transl_table=11
/product="CHEA HISTIDINE KINASE"
/protein_id="CAB50456.1"
/db_xref="GI:5458970"
/transl_table="MRSSELYKDIFREASINIMASHLTSLSOMIGGPIEMAEADVEIY
SRAEFKLIAERGVSKGTFVDFEGLSGLTILQPKHSALINISAVLGMGPGSMEI
LDMKRSALMEGNILISAVTDLISLIGEPYLSPPRAESLYDEKRLDRLRV
NSIVYFKSFYEDICVESYFYLVPPESFTLVKLEKQISEVEANDEN"
complement(10814..11452)
/note="PAB1333"
/codon_start=1
/transl_table=11
/product="CHEA HISTIDINE KINASE related protein"
/protein_id="CAB50457.1"
/db_xref="GI:5458971"
/transl_table="MDIVANTIEVYKNDIFAKSALLETINIGASHAATLSSEMGR
EVNISVPLDKVYKINPEIYVEDVAVVYIELGDFSSHAFFADYDLRFEDVIM
GNPPTTLELDMMVKSFMENGNILISAPANALISQILGITQTPALITDLPALID
LALADIGKFCDDTIILOTDIKISGVFEFEEHFLPKPEMDKRLKLEGLGLA"
complement(11436..13754)
/note="PAB1332"
/codon_start=1
/transl_table=11
/product="chemotaxis histidine kinase (chea)"
/protein_id="CAB50458.1"
/db_xref="GI:5458972"
/transl_table="MYTLEDMSOYIDEFADARDRIDSLSNLTIEKIVKGGSEEE
KABLINOIFRDHATLKRAATMGFKLSTFAHKMELFAIRNSQIELPEYVDVL
FLDALFAMVNNIEESGSDIDVSLFRADFLQKGAEKPEEKREVEANEE
EKPERVEGVKGVYHVKYFQKDAFLKGVRAFLISDIEEGEVYIWTBPGDVIIEG
NLDKVIKRVSSDLPKELEKVIYHPEVEKEIIEESEEEKGPEKEKYLIRYF
SPDAPLKGRSYLLODLESGIDIVOTNRGDIENGNIIEKGFPEVYRTBEDKSI
VOLIKKHDPVENYVTEAREKQEKKEKABEKRPQKQKRPYIEPRKISII
KIDVGHDLAMNVEELVITTKRLEOIGRLDRELFELSLTSLRLDLQDIBEMRV
LPIAEVENKFPYAVDLARKKKEVEFMEGADIEVDFTILEKLDALVHLRNMVD
HGLEPERVVKLGKPRVGRVELAKREKNHVVYIYVDDRGIDPEVKKKAIEKGLIT
PEBASISDEEALINILFLPGFSELAKREKTVDSRGVMDVYKVKMNGSISQVEVG
KGTTFILKPISMALIOALLIRYMDETVAIPNNILETIEVDPSTILKTGGKPVYLR
GLIIPVYMLHELGLPTPOVERPPATIIYHGAOKVAIGVDELHHRDIYIKSLKRLS
NVRGFRAGATIIDSGSVYLIIDGGLIGGYSG"
complement(13717..14823)
/note="PAB1331"
/codon_start=1
/transl_table=11
/product="protein-glutamate methyltransferase (chea)"

```

Query Match 8.18; Score 150.8; DB 1; Length 265118;
 Best Local Similarity 50.88; Pred. No. 1.7e-19;
 Matches 527; Conservative 0; Mismatches 487; Indels 24; Gaps 6;

```

QY 796 GAAGCAAGCTCAGGAGAAAGCTCTCTCCATCCAGATGCGGGAGGGAAC 855
DB 59002 GAACCTTATCTTTAAGATGCGTGGATATCTTGCCACAGGATGGGCTTAAT 58943
QY 856 CTCT-----GGCTGAGGACGGAGGCGGAGCGTGTGTTACCGGCACTGAT 909
DB 58942 CTGTTACGGGTAAACCTAGATGGAAGGTTAAGGATTAAGGAGATGAACGTT 58883
QY 910 TACGGCTTGACGTACGCGATGGCAAGACATTCCTCATCATGACCGCCACGAGATA 969

```

```

DB 58882 GAGAGCTTTCCGTTGGTGTATTACATAGCATTTTACAGCTCAAGATCCGACAACTCCGACT 58823
QY 970 GCGGACCTCTACCTCTTACAGCGCGAGCTGAAGACAGGTACCGGAATACAGCGGCCGATA 1029
DB 58822 GAGCTTTACGCTCTCAGGATGAGAAAGAGGACGTATCACAGACTTTAAACGCTGGA-- 58765
QY 1030 TTCAGAACCTCAAGACCTTCAGACCGGAGCACTTCCGTTCAAG--AGCAAAACCTC 1086
DB 58764 -TTAGGAGCTATTAATCACTTCAAAACCGAGCACTCAGGCTTACGCGTACGATGCGGAAG 58706
QY 1087 GAGATAGAGGGCTGATCTCAGCGCGAGGTAAAG--GAGAGGCGCCCGGTG 1140
DB 58705 GAAATAGATGCTCGATATCATGAAAGCGGTGACTTTAAGCTCGAAGAGAAATATCTGGA 58646
QY 1141 ATAGCTTCCTCCACGCGCGCGCGGAGGCGATGACGACACCGCTTCGCTACAGATG 1200
DB 58645 GTCCTGGAATTCACCGCTGCGCAAAAACAGCCCTACGTTATGCAATTTATGCAAGTTT 58586
QY 1201 CAGCTGATGGCGAAGGCGTACTACTGCTCTGCTGTAACCGCGCGCGGCGAGCGAC 1260
DB 58585 CAGGTTTATGTTGCGCAAGGCTTCGTTGAATCTTTTCAAAATCCAAAGGGAATGATGCT 58526
QY 1261 TATAGCAAGACTTCGCGCTCCGCGTCTGAGAGAGACTGCGTTGAGACTTTGAGAC 1320
DB 58525 TATGAGAGAGAGTTCGCG--GACATTAAGGGGTCACTACGAGAGCGGATATCAAGAC 58469
QY 1321 ATATATAGAGGATTCAGAGAGTCTTCAAGCTCGAACCGGACCGGAGAGGCGCTT 1380
DB 58468 CTATATGAGAGTATGTTACAGAGGCTTAAAGCCTTTGACTTATATGATGAGGAAGCCTT 58409
QY 1381 GGAATATACGGGATTAAGCTACGCGGCTTCATGACCACTGCGGCTTACTGACTGAGACGC 1440
DB 58408 GCGGTTACCGGAGATCATACGAGAGGCTTCATGACAAATGATGTTGGGACACAAAC 58349
QY 1441 CTCTTCAAGCAGAGAAATACGGAACGAGCATAGCTACTGCTACCAAGCTTACGCTTC 1500
DB 58348 AGATTCGAAAGCTGCGCTTACTACGAGCGCTCAATATCAACTGATGATGATCTTTGGAACG 58289
QY 1501 TCGGACATATGGGCTGCTGATACGAGCTGCGAGGTCATCGGGCCAAACCGCTTACAGAAC-- 1557
DB 58288 ACCGATATGCGCTATTACTTCTGCTCCAGATGAGATCGGCGGCACTTGAGATACCTTA 58229
QY 1558 GAGAACCTTCAAGAACCTCAGCCGCTGTCTACGCTCAGAACGTCGAAGCGCGGATATCTC 1617
DB 58228 GAAGGCTACTCGGAGAAAGAGCCATTTGAAGTACGCTCCAAAGTGAAGCTCGGCTCG 58169
QY 1618 CTATTCACCTCGCTTACGAGACTACCGCTTCGCTCGACCAAGCCTTATGTTTACAAAC 1677
DB 58168 ATATTCACCTCCACCGAGAGACTACCGCTTGTGCTCCGAGGCTTTGCAATTTTACATA 58109
QY 1678 GTGCTCAAGAGACATGGGCAAGGAGGCTCATATAGGATATTTCAAGGCGGCGCCACGCGC 1737
DB 58108 GCGTTCAAGATTTTGGGCAAAACCGTCGAGCTGGCGATTTTCCACAGTACAGATATGAC 58049
QY 1738 CACAGCGTCCGCGAAGCCCGAGGACAGGCGGAGCGGATACAGGCTCTTCTATAGATTTC 1797
DB 58048 TTGATGATGAGGTGAAGCCAAAGCATAGGTTAAGAGGCTCGAGCTCATAGTGGGGTGG 57989
QY 1798 TTCGAGCGCAAGTCAAG 1815
DB 57988 TTTGAGAGGTGCTGAAG 57971

```

RESULT 12
 AX041922/c
 LOCUS AX041922 265118 bp DNA linear PAT 23-NOV-2000
 DEFINITION Sequence 817 from Patent W00065062.
 ACCESSION AX041922
 VERSION AX041922.1 GI:11340687
 KEYWORDS
 SOURCE Pyrococcus abyssi.
 ORGANISM Pyrococcus abyssi
 Archaea; Euryarchaeota; Thermococci; Thermococcales;

REFERENCE
1 (bases 1 to 265118)

Thermococcaceae: Pyrococcus.

AUTHORS

Forrester, P., Thierry, J.C., Prieur, D., Dietrich, J., Lecompte, O.,
Querrelou, J., Weissenbach, J., Saurin, W., Helling, R., Flament, D.,
Raffin, J.P., Henneke, G., Gueguen, Y. and Rolland, J.L.
Genome sequence and polypeptides of pyrococcus _1(abissay), fragment
and uses thereof

JOURNAL

Patent: WO 0065067-A 817 02-NOV-2000;

CEMRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (FR) ; IREMER
INSTITUT FRANCAIS DE RECHERCHE POUR L'EXPLOITATI; ON DE LA MER (FR)

FEATURES

source

Location/Qualifiers

1..265118

/organism="Pyrococcus abyssi"

/db_xref="taxon:29292"

/note="Original length of seq 1: 1765118 (for 812
sequences) New sequences: seq 001, from 0.000.001 to
0.349.980, length: 349.980 seq 813, from 0.300.001 to
0.649.980, length: 349.980 seq 814, from 0.600.001 to
0.949.980, length: 349.980 seq 815, from 0.900.001 to
1.249.980, length: 349.980 seq 816, from 1.200.001 to
1.549.980, length: 349.980 seq 817, from 1.500.001 to
1.765.118, length: 265.118"

BASE COUNT 75226 a 61302 c 54005 g 74585 t
ORIGIN

Query Match 8.1% Score 150.8; DB 6; Length 265118;
Best Local Similarity 50.8%; Pred. No. 1.7e-19;
Matches 527; Conservative 0; Mismatches 487; Indels 24; Gaps 6;

796 GAAGCAAGCTCAGCGAAGAAAGTCTACTCTCTACCTCCAGATGCGGGGCTAAAC 855
Db GAACTTATCTTAAAGATGGGTGATATCTTCTCGCCAGATGGCGCTTAAAT 58943
QY 856 CTCF-----GGCTCTGGACGGGAGCCGAGCGTGTGTATACCGGACCACTGAT 909
Db CTGTTCAAGGTAACTTAAAGGTAAGGTTAAAGGTTAAAGGTAAGGTAAGGTT 58883
QY 910 TACGGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 969
Db GAGAGCTTTTCCGTTGGTGGTATATATATATATATATATATATATATATATAT 58882
QY 970 GCGAGCTTACCTCTACGACGCGGAGCTGAAACAGTACCAAGTACCAAGTACCAAG 1029
Db GAGCTTACCTCTACGACGCGGAGCTGAAACAGTACCAAGTACCAAGTACCAAG 58822
QY 1030 TTGGAAGCTCAAGACTTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 1086
Db TTGGAAGCTCAAGACTTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 58764
QY 1087 GAGATAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db GAGATAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 58705
QY 1141 ATAGTCTTCTCCAGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Db ATAGTCTTCTCCAGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 58645
QY 1201 CAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db CAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 58585
QY 1261 TATAGGAGAGCTGCGGCTCCGCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db TATAGGAGAGCTGCGGCTCCGCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAG 58525
QY 1321 ATATAGAGAGCTGCGGCTCCGCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1380
Db ATATAGAGAGCTGCGGCTCCGCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAG 58468
QY 1381 GGAATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db GGAATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 58408

QY 1441 CTCCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
Db AGATTAAGAGCTCCCTTAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCA 58348
QY 1501 TCGGACATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1557
Db ACCGATATGCGGATATATCTTCCCTCAGATCAGATCAGATCAGATCAGATCAG 58288
QY 1558 GAGAACTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1617
Db GAGGCTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 58228
QY 1618 CTATATCACTCCCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1677
Db ATATATCACTCCCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 58168
QY 1678 GTGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1737
Db GCGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 58108
QY 1738 CACAGCGTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1797
Db TTGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 58048
QY 1798 TTGAGGCGCAAGCTCAAG 1815
Db TTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 57987

RESULT 13
BSUB0017c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

1 (bases 1 to 217420)
Bacillus subtilis complete genome (section 17 of 21): from 3197001
to 3414420.
299120 AL009126
299120.1 GI:2635613
Bacillus subtilis.
Bacteria: Firmicutes: Bacillales: Bacillaceae: Bacillus.
1 (bases 1 to 217420)
Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.,
Azevedo, V., Bertoletti, M., Bessieres, P., Bolyard, A., Borchert, S.,
Boris, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C.,
Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V.,
Carter, N.M., Choi, S.K., Codani, J.J., Connerton, I.F., Cummings, N.J.,
Daniel, R.A., Denzot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,
Emmerson, P.T., Enllan, K.D., Erlington, J., Fabret, C., Ferrari, E.,
Fougeron, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A.,
Gallier, G., Giesse, G., Guy, B.J., Haga, K., Goffeau, A., Golligorsky, J.,
Haut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F.,
Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y.,
Klaer, B., Jones, L., Klein, C., Kobayashi, Y., Koetter, P.,
Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,
Lardinois, S., Lauber, D., Lazarevic, V., Lee, S.M., Levine, A., Liu, H.,
Masuda, S., Maue, C., Medigue, C., Medina, N., Mellado, R.P.,
Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M.,
Ogawa, K., Ogilvie, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M.,
Portetle, D., Porro, J.L., Prescott, A.M., Presecan, E., Pujic, P.,
Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M.,
Rivola, C., Rocha, E., Roche, B., Rose, M., Sadle, Y., Sato, T.,
Scanlan, E., Schleich, S., Schroeder, R., Scofield, F., Sekiguchi, J.,
Sekowska, A., Seror, S.J., Serron, P., Shin, B.S., Soldo, B.,
Sorokin, A., Taconi, E., Takagi, T., Takahashi, H., Takemaru, K.,
Tateuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A.,
Tosato, V., Uehyama, S., Vandendri, M., Vannier, F., Vassarotti, A.,
Viat, A., Wambolt, R., Wedler, E., Wedler, H., Weitzenecker, T.,
Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K.,
Yoshida, K., Yoshikawa, H.F., Zumbstein, E., Yoshikawa, H. and
Danchin, A.
The complete genome sequence of the gram-positive bacterium
Bacillus subtilis

JOURNAL	Nature 390 (6657), 249-256 (1997)
MEDLINE	98044033
PUBMED	9384377
REFERENCE	2 (bases 1 to 217420)
AUTHORS	Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.
TITLE	Submitted (18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur, Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr, danchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48
JOURNAL	Regulation de l'Expression Genetique, 28 rue du Docteur Roux, 75724
FEATURES	Location/Qualifiers
source	1. 217420
gene	/organism="Bacillus subtilis"
gene	/strain="168"
gene	/db_xref="taxon:1423"
gene	/complement(3. .1277)
gene	/gene="yulB"
gene	/complement(3. .1277)
gene	/gene="yulB"
gene	/function="unknown"
gene	/note="similar to L-rhamnose isomerase"
gene	/codon_start=1
gene	/transl_table=11
gene	/protein_id="CAB5107.1"
gene	/db_xref="GI:2635614"
gene	/db_xref="SPTREMBL:005264"
gene	/translation="MTIKANYDSAKQAYEKWGDIVBEALRQLEQVPISIHQWGDIE GEVNGELSGGIDVYGNPGKAPTEPEERLEKALSLPGKRVNLHAYETNRE AVERDELKPOHEFNWYMAKNIIGLDFNPTLSHEKAADGLTSHDPIREFWIR CIACRIGEFGEKELGPCPLINIWIPDGKIDIPSDRLTPRKRLKESJDFSEISQ HNDLSIESKLFSGSSEYVGSHEFYALVTHMKICLIDTGHHPHPEVSNKISSNL LYTDKIALHYSRPRVMSDHYVLDDELRIALAEIYRNHLEKVAIGDFPDMSIN AAVTIGTRNKRKLLYALLPNGLYKLOQEGERTERLALMEEFKTYPPGALWDSICE OMQVPEKAMLYIKEDQVLLKRRASPTV"
gene	/complement(1303. .1617)
gene	/gene="yulD"
gene	/complement(1303. .1617)
gene	/gene="yulD"
gene	/function="unknown"
gene	/note="similar to hypothetical proteins"
gene	/codon_start=1
gene	/transl_table=11
gene	/protein_id="CAB5108.1"
gene	/db_xref="GI:2635615"
gene	/db_xref="SPTREMBL:005263"
gene	/translation="MKRKASIMFVHODKTEEYQRRHDDIPEPAEALKAHGHYSIE LDEETGRLEFAYLEIEDEEKWRKMADEVCORWKSMAVPLKTNDSFPAVIDKEVY LD"
gene	/complement(1635. .3092)
gene	/gene="yulC"
gene	/complement(1635. .3092)
gene	/gene="yulC"
gene	/function="unknown"
gene	/note="similar to rhamnulokinase"
gene	/codon_start=1
gene	/transl_table=11
gene	/protein_id="CAB5109.1"
gene	/db_xref="GI:2635616"
gene	/db_xref="SPTREMBL:005262"
gene	/translation="MTYTAIDVGASGGRINWGLNBEKGLDIOEIHFRPANGFSQDGH LMDIDLKQILOGLOKVLGEGHCVGIDWAVYLLDEKGDRLRELAISYDRRT DRTIDLEHLSKAIYOKTIOFOPNTYOFEEDRELLKTKDKIMIPDLYCYL TGKAVETITVSTTOILNVTSGNDPELLAVLISQOAPLPEPGCEGKLGLENMP DYDLPACKWTVATHTHNASAVIAPGVNCSWAYISSGWSLITVENKTPITITLLEN NNTNERNANTIRPLKNIIGWVIOEVKQOLADYSFOOLAEBAKTEPPOFTNLND KRLPENMIKEIOHCROTROKIPRYAGELACCIISNLAIYIAIKELLETTERKI EOPHIIGGARNDPLNOLITADMSKAVYAGPIEATATGNLMOMIAKAEVKDIKEAO VVRNSPEIKFTPKRIDRSTTIOSFOOTVAKLSK"
gene	/complement(3097. .3873)
gene	/gene="yulB"
gene	/complement(3097. .3873)
gene	/gene="yulB"
gene	/function="unknown"
gene	/note="similar to transcriptional regulator (DeoR family)"
gene	/codon_start=1
gene	/transl_table=11
gene	/protein_id="CAB5110.1"
gene	/db_xref="GI:2635617"
gene	/db_xref="SWISS-PROT:005261"
gene	/translation="MTVAREQKIVITVMMRSSIRYSELDFSYTTEITRDLEKLE KEKLSRHGGAVSQKESIEHSEREITVYERKAIHBAKYKSGDRITLDAST TANYMAKILPDIELVITNSKRAIENKSNISVISTGILLLESLSAGPLAERSL EYHVNKTEPLSCGFDVHNGMSDNEMOALKRMERSDQITLMADSSKMGNRRSH IASLDVSRITDSDGPASVAKLEDKKVTAVPLSKRG"
gene	/complement(3930. .5999)
gene	/gene="yuxG"
gene	/complement(3930. .5999)
gene	/gene="yuxG"
gene	/function="unknown"
gene	/note="alternate gene name: yulA similar to sorbitol-6-phosphate 2-dehydrogenase"
gene	/codon_start=1
gene	/transl_table=11
gene	/protein_id="CAB5111.1"
gene	/db_xref="GI:2635618"
gene	/db_xref="SWISS-PROT:P40747"
gene	/translation="MTKIMDSERAOQKRGVELYRSNLIGSDRTVCMGGNTSM KTEKDPGRGEITVMYKSGSDLAEMKAKHNSGLTLDITRPLIKRDKQPDMDVYL SHCMIDSKHPRIETLHAFLEPYKVDHHPALITSCADNGKOIADITYGNRPV VPYVRPGETLSKMAIAGVANPAPLAVLEKRLVHGETSEYCTQITSIIOEAYQ INRINOHVEFGKRPOPLVEDRKQILAGIMVINGAASEKKMLSTDHDDVLEF VNSVQALSOIJAACPDHLYHTKRVPLIDWNPEDOVHKLADLKSEVEFTSEYQ AYTRNOODSDIFESAPRYILIGVWNTGKSYMSVSGALYRRAIYAKMGATL GSHVLAHNSYHVEVPLEYLTLEIAPPEERSRVALITGGAGIGSAACRRFAE GGHYIYADNLIEBQKTAGIENDAYCKGRAMVKNVTEEDYQKTFERRAALAYGID IVNNGSLATSPDEFSLKMNLMNVNIGTGFVLARAFKQMKHNGKSAVFGS KNSVYAGKNAISVAKLETHLARCILABGGEGFIRVNSVLPDAVLQCSALMGSSWR EERRAAYGIEPDLIEHYRKRKTLVNIYIPEDIAEIAEFASSKAKERTGCMITVDG VPAAFTFR"
gene	/complement(6109. .8125)
gene	/gene="tlpB"
gene	/complement(6109. .8124)
gene	/gene="tlpB"
gene	/complement(6137. .8125)
gene	/gene="tlpB"
gene	/codon_start=1
gene	/transl_table=11
gene	/product="methyl-accepting chemotaxis protein"
gene	/protein_id="CAB5112.1"
gene	/db_xref="GI:2635619"
gene	/db_xref="SWISS-PROT:P39217"
gene	/translation="MGKFTQIKRQPSISKPLIAPLAVILPVGVLAFYSOSANAL DRELISAKNGVFNELSTONKLEDKVADIVSEYVDDIILGKRTLLKEFKQYT TUNDGAIITASEDKKLYKYPDSGPKGFDPRGRWYQOAYAEKQAVFSEPYDEA TGDIIVATISKQLDGSGVIALDLNDEVLASRIKIGKEGFAITYGKKKTIHPTT KPGTSGDWTNQVYSKREGSEFEYEGEKKAFTNKLATGKLAGTGVSELDAS SPVLANVAILCVISVIGGILILYIIRATTKPRKIVSTSAKISGLDEVIDHSN EFGLOESFEMASLSRVYGIOTSEVNASSEELTSAAQTSKATHEITLAEQF SDGNEAOSEKLETSSNHLISOMNGISKVAQASSTIKRSIOSSEAGSEKLEIYHVG QAKTIQOSYOKAARVYKLETKSODITSTINYNGLADONTLAINALIAEAPAGVY KGRSVAAEYKRLAVOSADSAKEIBGLIDETVAIEISTISLMSQVNHREKGIQITDQ TASFKOITVETQIISGELONLNAVTEQLSAGQEVSSAVEDISAVAKKSSAGTODIA ASAEBOIASMEESSEETLANNAEELODITRKFKIES"
gene	/complement(8204. .10224)
gene	/gene="mcpA"
gene	/complement(8204. .8228)
gene	/gene="mcpA"
gene	/complement(8239. .10224)
gene	/gene="mcpA"
gene	/function="chemotaxis deficiency toward glucose and alpha-methyl-glucoside"
gene	/codon_start=1
gene	/transl_table=11
gene	/product="methyl-accepting chemotaxis protein"
gene	/protein_id="CAB5113.1"

/db_xref="GI:2635620"
 /db_xref="SWISS-PROT:P39214"
 /translation="MKKLIIDIKORSTIKLVEFLSILITIPVILAEVAVASSSL
 DROMGSALENOOLNEINTSIGEKNSADYEEMTLKEKYNKSNVSTREKSOIT
 SINKDVSITSDIKGHETRYPOLPMSGSDNPVRDMLKRAVANKGVITDPKITS
 TMTPTVIAOQTDGSGVIAINMTIENMLTKVNGTOSIAIMKDKVAVHPNE
 OSGTELDMLDKLSDKDPQYTMGDGDKKMFADNKLGMVIGMIDELTHEA
 OPLVHLALIVLAAIIGIIVMLIIRSTYTPKQLVGSKRISGDLTETIDIRSD
 ELGELGSPNNMASLSRLIHAIDSVDNVAASEELTASAOATSKAPHEHTLEOF
 SNGERKENENIETAEHIIYOMDGLTMAOASEVITDSVOSPEIASSEGCTVHOCTG
 OMNVIDKSVAEQVYRGLETKSDITNIIIRVINGINDORHLALNAIEARAGEY
 RPSVAEERKLAQVSDSAKEIEGLITIEVEIKINSLHFSQVNEVOTGLDITDK

Query Match 8.0%; Score 150.2; DB 1; Length 217420;
 Best Local Similarity 52.3%; Pred. No. 2.2e-19;
 Matches 357; Conservative 0; Mismatches 323; Indels 3; Gaps 1;

1135 CCGGTGATGCTTCGTCACAGGGGGCCGAGGCGCATGTACAGACCCGCTGCTAC 1194
 113504 CCCTTATCTTAAACATACACAGGGGGCTCCATATGATGATACGACATATTTTCAT 113445
 1195 GAGATGAGCTGATGAGGAGCAAGGCGTACTAGCTGCTGCTGGAACCGCGGCGAGC 1254
 113444 GAGTTTCAGGTGCTGGGGGGAAGATAGCGGCTCTTATATCAATCCGAGAGAGC 113385
 1255 GAGGCTATAGCGAAGACTTCCGCGCTCCGCTCGAGAGAGAGACTTGGAGACTTT 1314
 113384 CACGGCTACGGGAGGAGATTTTGAATGCGGCTACAGAGATTTATGGGGAAGATTA 113325
 1315 GAGACATATGAAAGGAGCATGAGAGTTCTCAAGCTGACACCGGAGCGGAGAGAG 1374
 113324 GAGAGATGATGAGAGCTGTGATGAGGCTATCAAGAGATCCGCTATTTGATCTTAAG 113265
 1375 CGCGTTGGAATTAACGGGCGATTAAGCTACGGGCTTCATGACCAACTGGGCTGACTAC 1434
 113264 CGGCTGAGTGTACGGGCGGAAAGCTACGAGGTTTATGACCAACTGATGCTGCGGAG 113205
 1435 AGGACCTCTTCAAGGAGAAATTAAGCGAGAGAGGCTTAAGTACTGCTACAGCTAC 1494
 113204 ACAGACCGCTTAAAGCTGCGCTACCGAGCTCGATTAATCAATTTGATCACTTTAC 113145
 1495 GCCTTCGAGACATAGGGCTCTGTGACAGCTGAGAGTATCGGGCAATTCGTTAGAG 1554
 113144 GGGGTAGATGATGCGGCTATTTCTTACAGACTGCGAGCTTGAAGCTGATGATTTGAG 113085
 1555 AAC---GAGAACTTCAGGAAGCTCAGCCGCTGTTCAGCTCAGAACGTGAAGGCGCG 1611
 113084 GACACAGAAAAGCTCTGGGACCGGCTCTCTTTAAATATAGCAGCAAAAGTGAAGACCG 113025
 1612 ATACTCCTTAATCCACTCGCTTGAGACTACCGCTGTCCGCTGACACAGAGCTTATGTTTC 1671
 113024 CTTTATATCTGATGATGAGGAGGAGGATGACCGATGCGCATGAGAGGCGGAGAGCTG 112965
 1672 TACACGTGCTCAAGGACATGGGCAAGAAAGCTTACATAGCGATTTTCAAGCGCGGCGC 1731
 112964 TTTATGCTCTGAAAAAATTTGGGCAAGAAACCAACTGTGTCCGTTTCCGAATGATGC 112905
 1732 CAGGCGCAACAGCTCCGCGGAAGCCCGAGGACACAGGCGTACAGGCTCTTCATA 1791
 112904 CACAAATTTATACAGCGACCGGACACCAAGACGCGGATCAAGGCTGATTTATATCAGC 112845
 1792 GAGTCTTCGAGGCGAAGCTCA 1814
 112844 TCATGCTTGTATCAACATCTCTA 112822

RESULT 14
 LOCUS AX006271 1971 bp DNA linear PAT 24-AUG-2000
 DEFINITION Sequence 1 from Patent W09903984.
 ACCESSION AX006271
 VERSION AX006271.1 GI:9929124
 KEYWORDS

SOURCE
 ORGANISM Bacillus subtilis.
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE 1 (bases 1 to 1971)
 AUTHORS Estell,D.A.
 TITLE Proteases from gram-positive organisms
 JOURNAL Patent: WO 9903984-A 1 28-JAN-1999;
 GENENOR INTERNATIONAL B.V (NL); ESTELL DAVID A (US)
 FEATURES
 SOURCE location/Qualifiers
 1..1971
 /organism="Bacillus subtilis"
 /db_xref="taxon:1423"

BASE COUNT 558 a 431 c 524 g 458 t
 ORIGIN

Query Match 8.0%; Score 149.8; DB 6; Length 1971;
 Best Local Similarity 52.3%; Pred. No. 2.2e-19;
 Matches 356; Conservative 0; Mismatches 322; Indels 3; Gaps 1;

1135 CCGGTGATGCTTCGTCACAGGGGGCCGAGGCGCATGTACAGACCCGCTGCTAC 1194
 1291 CCCTTATCTTAAACATACACAGGGGGCTCCCATATGATGATACGACATATTTTCAT 1350
 1195 GAGATGAGCTGATGAGGAGCAAGGCGTACTACTGCTGCTGGAACCCGCGGCGAGC 1254
 1351 GAGTTTCAGGTGCTGGGGGGAAGATAGCGGCTGTTATATCAATCCGAGAGAGAGC 1410
 1255 GAGGCTATAGCGAAGACTTCCGCGCTCCGCTCGAGAGAGAGCTGCTTGGAGACTTT 1314
 1411 CACGGCTACGGGAGGAGATTTTGAATGCGGCTACAGAGATTTATGGGGAAGATTA 1470
 1315 GAGACATATGAAAGGAGCATGAGAGATTTCTCAAGCTCAGAACCGGAGCGGAGAGAG 1374
 1471 GAGAGATGATGAGAGCTGTGATGAGGCTATCAAGAGATCCGATATTTGATCTTAAG 1530
 1375 CGCGTTGGAATTAACGGGCGATTAAGCTACGGGCTTCATGACCAACTGGGCTTACTAC 1434
 1531 CGGCTGAGTGTACAGGGCGGAGAGCTACGAGGTTTATGACCAACTGATGCTGCGGAG 1590
 1435 AGGACCTCTTCAAGGAGGAGATTAAGGAGAGAGGATTAAGTACTGCTACAGCTAC 1494
 1591 ACAGACCGCTTAAAGCTGCGCTACCGAGCTCGATTAATCAATTTGATCACTTTAC 1650
 1495 GCCTTCGAGACATAGGGCTCTGTGACAGCTGAGAGTATCGGGCAATTCGTTAGAG 1554
 1651 GGGGTAGATGATGCGGCTATTTCTTACAGACTGCGAGCTTGAAGCTGATGATTTGAG 1710
 1555 AAC---GAGAACTTCAGGAAGCTCAGCCGCTGTTCAGCTCAGAACGTGAAGGCGCG 1611
 1711 GACACAGAAAAGCTCTGGGACCGGCTCTCTTTAAATATAGCAGCAAAAGTGAAGACCG 1770
 1612 ATACTCCTTAATCCACTCGCTTGAGACTACCGCTGTCCGCTGACAGAGCTTATGTTTC 1671
 1771 CTTTATATCTGATGATGAGGAGGAGGATGACCGATGCGCATGAGAGGCGGAGAGCTG 1830
 1672 TACACGTGCTCAAGGACATGGGCAAGAAAGCTTACATAGCGATTTTCAAGCGCGGCGC 1731
 1831 TTTATGCTCTGAAAAAATTTGGGCAAGAAACCAACTGTGTCCGTTTCCGAATGATGC 1890
 1732 CAGGCGCAACAGCTCCGCGGAAGCCCGAGGACACAGGCGTACAGGCTCTTCATA 1791
 1891 CACAAATTTATACAGCGACCGGACACCAAGACGCGGATCAAGGCTGATTTATATCAGC 1950
 1792 GAGTCTTCGAGGCGAAGCTC 1812
 1951 TCATGCTTGTATCAACATCTCTC 1971

RESULT 15
 LOCUS AX431809 1958 bp DNA linear PAT 28-JUN-2002
 DEFINITION Sequence 224 from Patent W00229113.
 ACCESSION AX431809

